

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
14 March 2002 (14.03.2002)

PCT

(10) International Publication Number  
**WO 02/20815 A2**

(51) International Patent Classification<sup>7</sup>: C12P 7/02, C12N 9/10, 9/88 // (C12P 7/02, C12R 1:26)

Rockland, DE 19732 (US). KOFFAS, Mattheos [GR/US]; Apartment D, 1013 Cloister Road, Wilmington, DE 19809 (US). WANG, Siquan [US/US]; 4 Collins Drive, Wilmington, DE 19803 (US).

(21) International Application Number: PCT/US01/26853

(74) Agent: FELTHAM, S., Neil; E.I. Dupont De Nemours and Company, Legal Patent Records Center, 1007 Market Street, Wilmington, DE 19898 (US).

(22) International Filing Date: 28 August 2001 (28.08.2001)

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(25) Filing Language: English

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(26) Publication Language: English

(30) Priority Data:

60/229,907 1 September 2000 (01.09.2000) US  
60/229,858 1 September 2000 (01.09.2000) US

(71) Applicant (*for all designated States except US*): E.L. DUPONT DE NEMOURS AND COMPANY [US/US]; Legal Patent Records Center, 1007 Market Street, Wilmington, DE 19898 (US).

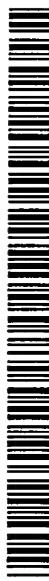
[Continued on next page]

(72) Inventors; and

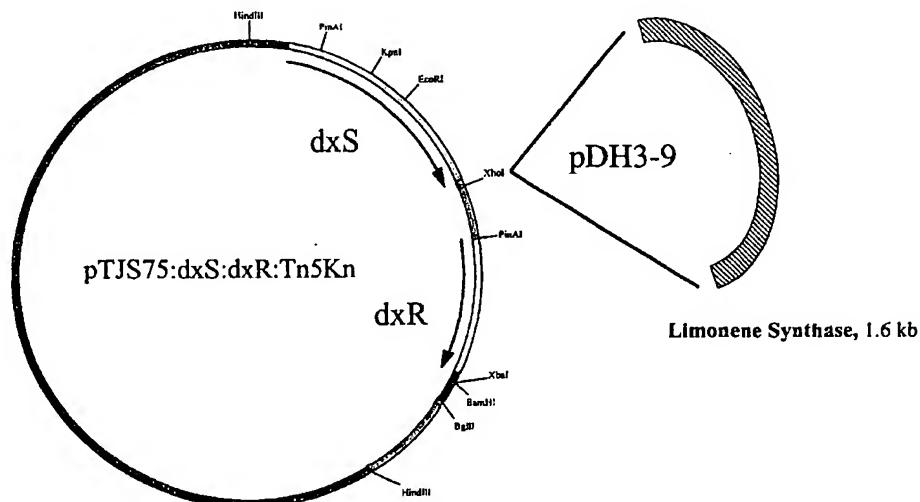
(75) Inventors/Applicants (*for US only*): DICOSIMO, Deana, J. [US/US]; 42 Millstone Lane, P.O. Box 330,

(54) Title: PRODUCTION OF CYCLIC TERPENOIDS

## Expression vector and construct



WO 02/20815 A2



(57) Abstract: A methanotrophic bacterium has been genetically engineered to produce cyclic terpenoids from geranyl pyrophosphate.



**Published:**

— without international search report and to be republished upon receipt of that report

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

TITLE

## PRODUCTION OF CYCLIC TERPENOIDS

This application claims the benefit of U.S. Provisional Application No. 60/229,907, filed September 1, 2000 and the benefit of U.S. 5 Provisional Application No. 60/229,858 filed September 1, 2000.

FIELD OF THE INVENTION

This invention is in the field of microbiology. More specifically, this invention pertains to methods for the production of cyclic terpenoid compounds in microbial hosts that metabolize single carbon substrates as 10 a sole carbon source.

BACKGROUND OF THE INVENTION

Monoterpene have value in the flavor and fragrance industries, as components in industrial solvents and in the pharmaceutical industry where selected compounds have shown promise as both chemopreventive 15 and chemotherapeutic agents for solid tumors.

Although found in a wide range of organisms, including bacteria, fungi, algae, insects and even higher animals such as alligators and beavers, monoterpene are most widely produced by terrestrial plants such as components of flower scents, essential oils, and turpentine. One 20 of the most common sources of the monoterpene are the herbaceous plant and conifer turpentine. The pinene regioisomers ( $\alpha$ -pinene,  $\beta$ -pinene) are 2 principal monoterpene of turpentine as they serve as large volume aroma chemicals. Other essential oils (from orange, lime, lemon, and peppermint) are valued in flavoring and perfumery. The 25 cyclization of linear terpenoid compounds to form cyclic derivatives may generate diverse aromatic structures with differing functionality.

At present the monoterpene may be obtained either by extraction from natural sources or by chemical synthesis. Both processes are time consuming and expensive. Although small scale production of selected 30 monoterpene has been demonstrated in microbial hosts, a facile method for the production of monoterpene on an industrial scale has yet to be reported. For example some monoterpene synthases have been successfully cloned and expressed in *Escherichia coli*. Specifically, limonene synthase, which catalyzes the cyclization of geranyl diphosphate 35 to yield the olefin 4(S)-limonene in *Perilla frutescens* has been cloned into *Escherichia coli* and functionally expressed (Yuba et al. *Arch Biochem Biophys* 332:280-287, (1996)). Reports of microbial expression however

have been limited to microbe traditionally used for fermentative production were grown on complex carbon substrates.

There are a number of microorganisms that utilize single carbon substrates as sole energy sources. These organisms are referred to as

- 5 methylotrophs and herein as "C1 metabolizers". These organisms are characterized by the ability to use carbon substrates lacking carbon to carbon bonds as a sole source of energy and biomass. A subset of methylotrophs are the methanotrophs which have the unique ability to utilize methane as a sole energy source. Although a large number of these
- 10 organisms are known, few of these microbes have been successfully harnessed to industrial processes for the synthesis of materials. Although single carbon substrates are cost effective energy sources, difficulty in genetic manipulation of these microorganisms as well as a dearth of information about their genetic machinery has limited their use primarily to
- 15 the synthesis of native products. For example the commercial applications of biotransformation of methane have historically fallen broadly into three categories: 1) Production of single cell protein, (Sharpe D. H. BioProtein Manufacture 1989. Ellis Horwood series in applied science and industrial technology. New York: Halstead Press.) (Villadsen, John, *Recent Trends*
- 20 *Chem. React. Eng.*, [Proc. Int. Chem. React. Eng. Conf.], 2nd (1987), Volume 2, 320-33. Editor(s): Kulkarni, B. D.; Mashelkar, R. A.; Sharma, M. M. Publisher: Wiley East., New Delhi, India; Naguib, M., Proc. OAPEC Symp. Petroprotein, [Pap.] (1980), Meeting Date 1979, 253-77 Publisher: Organ. Arab Pet. Exporting Countries, Kuwait, Kuwait.); 2) epoxidation of
- 25 alkenes for production of chemicals (US 4348476); and 3) biodegradation of chlorinated pollutants (Tsien et al., *Gas, Oil, Coal, Environ. Biotechnol.* 2, [Pap. Int. IGT Symp. Gas, Oil, Coal, Environ. Biotechnol.], 2nd (1990), 83-104. Editor(s): Akin, Cavit; Smith, Jared. Publisher: Inst. Gas Technol., Chicago, IL.; WO 9633821; Merkley et al., *Biorem. Recalcitrant Org.*, [Pap.
- 30 Int. In Situ On-Site Bioreclam. Symp.], 3rd (1995), 165-74. Editor(s): Hinchee, Robert E; Anderson, Daniel B.; Hoeppel, Ronald E. Publisher: Battelle Press, Columbus, OH: Meyer et al., *Microb. Releases* (1993), 2(1), 11-22). Even here, the commercial success of the methane bio-transformation has been limited to epoxidation of alkenes due to low
- 35 product yields, toxicity of products and the large amount of cell mass required to generate product associated with the process.

One of the most common classes of single carbon metabolizers are the methanotrophs. Methanotrophic bacteria are defined by their ability to use methane as a sole source of carbon and energy. Methane monooxygenase is the enzyme required for the primary step in methane activation and the product of this reaction is methanol (Murrell et al., *Arch. Microbiol.* (2000), 173(5-6), 325-332). This reaction occurs at ambient temperature and pressures whereas chemical transformation of methane to methanol requires temperatures of hundreds of degrees and high pressure (Grigoryan, E. A., *Kinet. Catal.* (1999), 40(3), 350-363; 5 WO 2000007718; US 5,750,821 ). It is this ability to transform methane under ambient conditions along with the abundance of methane that makes the biotransformation of methane a potentially unique and valuable process.

Many methanotrophs contain an inherent isoprenoid pathway which 15 enables these organisms to synthesize other non-endogenous isoprenoid compounds. Since methanotrophs can use one carbon substrate (methane or methanol) as an energy source, it is possible to produce monoterpenes at low cost. Furthermore, during the fermentation, volatile compounds can be easily removed as methane is passed through 20 fermentation media. It is also advantageous to produce via bio-route since many of monoterpenes have chirality and it is difficult to control the synthesis and purification of specific chirally active compound in chemical synthesis.

A need exists therefore for a method of production of highly 25 valuable monoterpenes from an inexpensive feedstock. Applicants have solved the stated problem by providing a C1 metabolizing microorganism having transformed with a gene encoding a cyclic terpene synthase, having the ability to produce to a variety of monoterpenes.

#### SUMMARY OF THE INVENTION

30 The invention provides a method for the production of a monoterpene comprising:

- a) providing a transformed C1 metabolizing host cell comprising:
  - (i) suitable levels of geranyl pyrophosphate; and
  - (ii) at least one isolated nucleic acid molecule encoding a 35 cyclic terpene synthase under the control of suitable regulatory sequences;

(b) contacting the host cell of step (a) under suitable growth conditions with an effective amount of a C1 carbon substrate whereby a monoterpenes compound is produced.

Preferred single carbon substrates of the present invention include  
5 but are not limited to methane, methanol, formaldehyde, formic acid, methylated amines, methylated thiols, and carbon dioxide.

Preferred C1 metabolizers or facultative methylotrophs where obligate methanotrophic bacteria are most preferred. Most preferred C1  
10 metabolizers are those obligate methanotrophs comprising a functional Embden-Meyerhof carbon pathway, said pathway comprising a gene encoding a pyrophosphate dependent phosphofructokinase enzyme.

Preferred cyclic terpene synthases of the invention include but are not limited to limonene synthase, pinene synthase, bornyl synthase, phellandrene synthase, cineole synthase, and sabinene synthase.

15 In an alternate embodiment the invention provides for the expression of upper pathway isoprenoid genes for the downstream production of monoterpenes, the upper pathway isoprenoid genes selected from the group consisting of D-1-deoxyxylulose-5-phosphate synthase (DXS); D-1-deoxyxylulose-5-phosphate reductoisomerase (DXR); 2C-methyl-d-erythritol cytidyltransferase (IspD), 4-diphosphocytidyl-2-C-methylerythritol kinase (IspE), 2C-methyl-d-erythritol 2,4-cyclodiphosphate synthase (IspF), CTP synthase (IspA) and Geranyltranstransferase (PyrG).

#### BRIEF DESCRIPTION OF THE DRAWINGS,

#### SEQUENCE DESCRIPTIONS, AND BIOLOGICAL DEPOSITS

Figure 1 shows the map of pTJS75:dxs:dxr:Tn5Kn plasmid containing truncated limonene synthase gene.

Figure 2 shows the gas chromatography analysis of limonene produced in *Methylomonas 16a* culture.

30 Figure 3 shows the examples of monoterpenes derived from geranyl diphosphate.

Figure 4 illustrates the upper isoprenoid pathway

Figure 5 shows the growth of *Methylomonas 16a* compared to the growth of *Methylococcus capsulatus* under identical growth conditions.

35 Figure 6 is a Schematic of Entner-Douderoff and Embden-Meyerhoff pathways in *Methylomonas 16a* showing microarray expression results numerically ranked in order of decreasing expression level.

The invention can be more fully understood from the following detailed description and the accompanying sequence descriptions, which form a part of this application.

The following sequences comply with 37 C.F.R. 1.821-1.825

- 5 ("Requirements for Patent Applications Containing Nucleotide Sequences and/or Amino Acid Sequence Disclosures - the Sequence Rules") and are consistent with World Intellectual Property Organization (WIPO) Standard ST.25 (1998) and the sequence listing requirements of the EPO and PCT (Rules 5.2 and 49.5(a-bis), and Section 208 and Annex C of the
- 10 Administrative Instructions). The symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. §1.822.

SEQ ID NO:1-4 are primer sequences.

SEQ ID NO:5 is the nucleotide sequence of plasmid

- 15 pTJS75:dxS:dxR:Tn5Kn.

SEQ ID NO:6 is the nucleotide sequence of limonene synthase gene from *Mentha spicata* with 57 amino acid sequences deleted from N-terminal.

- 20 SEQ ID NO:7 is deduced amino acid sequence of limonene synthase gene used in SEQ ID NO:6.

Applicants made the following biological deposits under the terms of the Budapest Treaty on the International Recognition of the Deposit of Micro-organisms for the Purposes of Patent Procedure:

Depositor Identification Reference	International Depository Designation	Date of Deposit
<i>Methylomonas</i> 16a	ATCC PTA 2402	August 21 2000

25

#### DETAILED DESCRIPTION OF THE INVENTION

The present invention describes a method for the synthesis of monoterpenes in a recombinant methylotrophic or methanotrophic host. Monoterpenes are used in flavors and fragrances, coatings and nutrition and health applications.

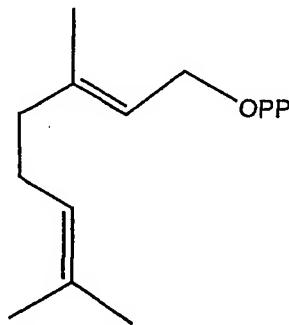
30 The following definitions may be used for interpretation of the claims and the specification.

"Open reading frame" is abbreviated ORF.

"Polymerase chain reaction" is abbreviated PCR.

The term "isoprenoid" or "terpenoid" refers to the compounds or any molecule derived from the isoprenoid pathway including 10 carbon terpenoids (monoterpene) and their derivatives, such as limonene, pinene, sabinene,  $\beta$ -phellandrene, borneol, carotenoids and xanthophylls.

- 5      The term "isoprene subunit" refers to a basic 5 carbon unit of isopentenyl diphosphate that further condenses to form a terpenoid.
- The term "cyclic terpene synthase" refers to an enzyme capable of using geranyl pyrophosphate as a substrate to produce a cyclic terpenoid compound.
- 10     The term "monoterpene" refers to any 10 carbon compound derived from geranyl pyrophosphate or its derivatives built upon 2 isoprene subunits (see Figure 3 for example).
- The term "cyclic monoterpene" refers to a cyclic terpenoid derived from geranyl pyrophosphate having 10 carbon atoms.
- 15     The term "limonene synthase" refers to enzyme that catalyzes the conversion of geranyl pyrophosphate to (-)-Limonene.
- The term "pinene synthase" refers to the enzyme that catalyzes the conversion of geranyl pyrophosphate to pinene.
- 20     The term "bornyl synthase" refers to the enzyme that catalyzes the conversion of geranyl pyrophosphate to borneol.
- The term "phellandrene synthase" refers to the enzyme that catalyzes the conversion of geranyl pyrophosphate to  $\beta$ -phellandrene.
- The term "cineole synthase" refers to the enzyme that catalyzes the conversion of geranyl pyrophosphate to cineole.
- 25     The term "sabinene synthase" refers to the enzyme that catalyzes the conversion of geranyl pyrophosphate to sabinene.
- The term "geranyl diphosphate" and "geranyl pyrophosphate" will be used interchangeably and will refer to a compound having the general formula



The term "Dxs" refers to the 1-deoxyxylulose-5-phosphate synthase enzyme encoded by the *dxs* gene.

The term "Dxr" refers to the 1-deoxyxylulose-5-phosphate reductoisomerase enzyme encoded by the *dxr* gene.

5 The term "YgbP" or "IspD" refers to the 2C-methyl-D-erythritol cytidyltransferase enzyme encoded by the *ygbP* or *ispD* gene. The names of the gene, *ygbP* or *ispD*, are used interchangeably in this application. The names of gene product, YgbP or IspD are used interchangeably in this application.

10 The term "YchB" or "IspE" refers to the 4-diphosphocytidyl-2-C-methylerythritol kinase enzyme encoded by the *ychB* or *ispE* gene. The names of the gene, *ychB* or *ispE*, are used interchangeably in this application. The names of gene product, YchB or IspE are used interchangeably in this application.

15 The term "YgbB" or "IspF" refers to the 2C-methyl-d-erythritol 2,4-cyclodiphosphate synthase enzyme encoded by the *ygbB* or *ispF* gene. The names of the gene, *ygbB* or *ispF*, are used interchangeably in this application. The names of gene product, YgbB or IspF are used interchangeably in this application.

20 The term "PyrG" refers to a CTP synthase enzyme encoded by the *pyrG* gene.

The term "IspA" refers to Geranyltransferase or farnesyl diphosphate synthase enzyme as one of prenyl transferase family encoded by *ispA* gene.

25 The term "LytB" refers to protein having a role in the formation of dimethylallyl-pyrophosphate in the isoprenoiod pathway and which is encoded by *lytB* gene.

The term "upper pathway isoprene genes" refers to any of the following genes and gene products associated with the isoprenoid biosynthetic pathway including the *dxs* gene (encoding 1-deoxyxylulose-5-phosphate synthase), the *dxr* gene (encoding 1-deoxyxylulose-5-phosphate reductoisomerase), the "ispD" gene (encoding the 2C-methyl-D-erythritol cytidyltransferase enzyme); the "ispE" gene (encoding the 4-diphosphocytidyl-2-C-methylerythritol kinase); the "ispF" gene (encoding a 2C-methyl-d-erythritol 2,4-cyclodiphosphate synthase); the "pyrG" gene (encoding a CTP synthase); the "ispA" gene (encoding geranyltransferase or farnesyl diphosphate synthase), and the "lytB" gene.

The term "single carbon substrate" refers to a carbon substrate useful as a microbial feedstock being devoid of carbon to carbon bonds.

The term "C1 metabolizer" refers to a microorganism that has the ability to use a single carbon substrate as a sole source of energy and

- 5 biomass. C1 metabolizers will typically be methylotrophs and/or methanotrophs.

The term "methylotroph" means an organism capable of oxidizing organic compounds which do not contain carbon-carbon bonds. Where the methylotroph is able to oxidize CH<sub>4</sub>, the methylotroph is also a

- 10 methanotroph.

The term "methanotroph" means a prokaryote capable of utilizing methane as a substrate. Complete oxidation of methane to carbon dioxide occurs by aerobic degradation pathways. Typical examples of methanotrophs useful in the present invention include but are not limited to

15 the genera *Methylomonas*, *Methylobacter* *Methylococcus*, and *Methylosinus*.

The term "*Methylomonas 16a*" and "*Methylomonas 16a sp.*" are used interchangeably and refer to the *Methylomonas* strain used in the present invention.

- 20 The term "Embden-Meyerhof pathway" refers to the series of biochemical reactions for conversion of hexoses such as glucose and fructose to important cellular 3 carbon intermediates such as glyceraldehyde 3 phosphate, dihydroxyacetone phosphate, phosphoenol pyruvate and pyruvate. These reactions typically proceed with net yield of
- 25 biochemically useful energy in the form of ATP. The key enzymes unique to the Embden-Meyerof pathway are the phosphofructokinase and fructose 1,6 bisphosphate aldolase.

- 30 The term "Entner-Douderoff pathway" refers to a series of biochemical reactions for conversion of hexoses suchas as glucose or fructose to important 3 carbon cellular intermediates pyruvate and glyceraldehyde 3 phosphate without any net production of biochemically useful energy. The key enzymes unique to the Entner-Douderoff pathway are the 6 phosphogluconate dehydratase and the ketodeoxyphosphogluconate aldolase.

- 35 The term "high growth methanotrophic bacterial strain" refers to a bacterium capable of growth with methane or methanol as sole carbon and energy source which possess a functional Embden-Meyero carbon flux

pathway resulting in a high rate of growth and yield of cell mass per gram of C1 substrate metabolized. The specific "high growth methanotrophic bacterial strain" described herein is referred to as "*Methyloimonas 16a*" or "16a", which terms are used interchangeably.

- 5        As used herein, an "isolated nucleic acid fragment" is a polymer of RNA or DNA that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. An isolated nucleic acid fragment in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA.
- 10      "Gene" refers to a nucleic acid fragment that is capable of being expressed as a specific protein, including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. "Native gene" refers to a gene as found in nature with its own regulatory sequences. "Chimeric gene" refers to any gene that is
- 15      15 not a native gene, comprising regulatory and coding sequences that are not found together in nature. Accordingly, a chimeric gene may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found
- 20      20 in nature. "Endogenous gene" refers to a native gene in its natural location in the genome of an organism. A "foreign" gene refers to a gene not normally found in the host organism, but that is introduced into the host organism by gene transfer. Foreign genes can comprise native genes inserted into a non-native organism, or chimeric genes. A
- 25      25 "transgene" is a gene that has been introduced into the genome by a transformation procedure.
- "Coding sequence" refers to a DNA sequence that codes for a specific amino acid sequence. "Suitable regulatory sequences" refer to nucleotide sequences located upstream (5' non-coding sequences), within,
- 30      30 or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include promoters, translation leader sequences, introns, polyadenylation recognition sequences, RNA processing site, effector binding site and stem-loop structure.
- "Promoter" refers to a DNA sequence capable of controlling the expression of a coding sequence or functional RNA. In general, a coding

sequence is located 3' to a promoter sequence. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic DNA segments. It is understood by those skilled in the

- 5 art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental or physiological conditions. Promoters which cause a gene to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". It is further recognized  
10 that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of different lengths may have identical promoter activity.

The term "operably linked" refers to the association of nucleic acid sequences on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence when it is capable of affecting the expression of that coding sequence (i.e., that the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation.

- 20 The term "expression", as used herein, refers to the transcription and stable accumulation of sense (mRNA) or antisense RNA derived from the nucleic acid fragment of the invention. Expression may also refer to translation of mRNA into a polypeptide.

- 25 "Transformation" refers to the transfer of a nucleic acid fragment into the genome of a host organism, resulting in genetically stable inheritance. Host organisms containing the transformed nucleic acid fragments are referred to as "transgenic" or "recombinant" or "transformed" organisms.

- 30 The terms "plasmid", "vector" and "cassette" refer to an extra chromosomal element often carrying genes which are not part of the central metabolism of the cell, and usually in the form of circular double-stranded DNA fragments. Such elements may be autonomously replicating sequences, genome integrating sequences, phage or nucleotide sequences, linear or circular, of a single- or double-stranded  
35 DNA or RNA, derived from any source, in which a number of nucleotide sequences have been joined or recombined into a unique construction which is capable of introducing a promoter fragment and DNA sequence

for a selected gene product along with appropriate 3' untranslated sequence into a cell. "Transformation cassette" refers to a specific vector containing a foreign gene and having elements in addition to the foreign gene that facilitates transformation of a particular host cell. "Expression cassette" refers to a specific vector containing a foreign gene and having elements in addition to the foreign gene that allow for enhanced expression of that gene in a foreign host.

Standard recombinant DNA and molecular cloning techniques used here are well known in the art and are described by Sambrook, J., Fritsch, E. F. and Maniatis, T., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989) (hereinafter "Maniatis"); and by Silhavy, T. J., Bennan, M. L. and Enquist, L. W., Experiments with Gene Fusions, Cold Spring Harbor Laboratory Cold Press Spring Harbor, NY (1984); and by Ausubel, F. M. et al., Current Protocols in Molecular Biology, published by Greene Publishing Assoc. and Wiley-Interscience (1987).

The present invention provides a method for the synthesis of monoterpenes in a methylotrophic or methanotrophic microbial host. Typical monoterpenes of the invention are those that are derived from geranyl pyrophosphate and contain ten carbon atoms. Typically the hosts have the elements of the isoprenoid pathway that will result in the production of geranyl pyrophosphate. The microbial host will also comprise a gene encoding a synthase, which is capable of using geranyl pyrophosphate as a substrate to produce a monoterpene.

#### 25 Identification and Isolation of C1 Metabolizing Microorganisms

The present invention provides for the expression of cyclic terpene synthases in microorganisms which are able to use single carbon substrates as a sole energy source. Such microorganisms are referred to herein as C1 metabolizers. The host microorganism may be any C1 metabolizer which has the ability to synthesize geranyl diphosphate (GPP), the precursor for many of the monoterpenes.

Many C1 metabolizing microorganisms are known in the art and are able to use a variety of single carbon substrates. Single carbon substrates useful in the present invention include but are not limited to methane, methanol, formaldehyde, formic acid, methylated amines (e.g. mono, di- and tri-methyle amine), methylated thiols, and carbon dioxide.

All C1 metabolizing microorganisms are generally classed as methylotrophs. Methylotrophs may be defined as any organism capable of oxidizing organic compounds which do not contain carbon-carbon bonds.

A subset of methylotrophs are the methanotrophs which have the distinctive ability to oxidize methane. Facultative methylotrophs have the ability to oxidize organic compounds which do not contain carbon-carbon bonds, but may also use other carbon substrates such as sugars and complex carbohydrates for energy and biomass. Obligate methylotrophs are those organisms which are limited to the use of organic compounds which do not contain carbon-carbon bonds for the generation of energy and obligate methanotrophs are those obligate methylotrophs that have the ability to oxidize methane.

Facultative methylotrophic bacteria are found in many environments, but are isolated most commonly from soil, landfill and waste treatment sites. Many facultative methylotrophs are members of the  $\beta$ , and  $\gamma$  subgroups of the Proteobacteria (Hanson et al., *Microb. Growth C1 Compounds.*, [Int. Symp.], 7th (1993), 285-302. Editor(s): Murrell, J. Collin; Kelly, Don P. Publisher: Intercept, Andover, UK; Madigan et al., Brock Biology of Microorganisms, 8th edition, Prentice Hall, UpperSaddle River, NJ (1997)).

Facultative methylotrophic bacteria suitable in the present invention include but are not limited to, *Methylophilus*, *Methylobacillus*, *Methylobacterium*, *Hyphomicrobium*, *Xanthobacter*, *Bacillus*, *Paracoccus*, *Nocardia*, *Arthrobacter*, *Rhodopseudomonas*, and *Pseudomonas*.

The ability to utilize single carbon substrates is not limited to bacteria, but also extends to yeasts and fungi. For example a variety of yeast genera are able to use single carbon substrates in addition to more complex materials as energy sources. Specific methylotrophic yeasts useful in the present invention include but are not limited to *Candida*, *Hansenula*, *Pichia*, *Torulopsis*, and *Rhodotorula*.

Those methylotrophs having the additional ability to utilize methane are referred to as methanotrophs. Of interest in the present invention are those obligate methanotrophs which are methane utilizers but which are obliged to use organic compounds lacking carbon-carbon bonds.

Exemplary of these organisms are included in, but not limited to the genera *Methylomonas*, *Methylobacter*, *Methylcoccus*, *Methylosinus*, *Methylcytis*, *Methylococcus*, and *Methanomonas*.

Of particular interest in the present invention are high growth obligate methanotrophs having an energetically favorable carbon flux pathway. For example Applicants have discovered a specific strain of methanotroph having several pathway features which make it particularly useful for carbon flux manipulation. This type of strain has served as the host in present application and is known as *Methyloimonas 16a* (ATCC PTA 2402).

The present strain contains several anomalies in the carbon utilization pathway. For example, based on genome sequence data, the strain is shown to contain genes for two pathways of hexose metabolism. The Entner-Douderoff Pathway which utilizes the keto-deoxy phosphogluconate aldolase enzyme is present in the strain. It is generally well accepted that this is the operative pathway in obligate methanotrophs. Also present, however, is the Embden-Meyerhoff Pathway which utilizes the Fructose bisphosphate aldolase enzyme. It is well known that this pathway is either not present or not operative in obligate methanotrophs. Energetically, the latter pathway is most favorable and allows greater yield of biologically useful energy and ultimately production of cell mass and other cell mass-dependent products in *Methyloimonas 16a*. The activity of this pathway in the present 16a strain has been confirmed through microarray data and biochemical evidence measuring the reduction of ATP. Although the 16a strain has been shown to possess both the Embden-Meyerhoff and the Entner-Douderoff pathway enzymes the data suggests that the Embden-Meyerhoff pathway enzymes are more strongly expressed than the Entner-Douderoff pathway enzymes. This result is surprising and counter to existing beliefs on the glycolytic metabolism of methanotrophic bacteria. Applicants have discovered other methanotrophic bacteria having this characteristic, including for example, *Methyloimonas clara* and *Methylosinus sporium*.

A particularly novel and useful feature of the Embden-Meyerhoff pathway in strain 16a is that the key phosphofructokinase step is pyrophosphate dependent instead of ATP dependent. This feature adds to the energy yield of the pathway by using pyrophosphate instead of ATP. Because of its significance in providing an energetic advantage to the strain this gene in the carbon flux pathway is considered diagnostic for the present strain.

In methanotrophic bacteria methane is converted to biomolecules via a cyclic set of reaction known as the ribulose monophosphate pathway or RuMP cycle. This pathway is comprised of three phases, each phases being a series of enzymatic steps (Figure 3). The first step is "fixation" or incorporation of C-1 (formaldehyde) into a pentose to form a hexose or six carbon sugar. This occurs via a condensation reaction between a 5 carbon sugar (pentose) and formaldehyde and is catalyzed by hexulose monophosphate synthase. The second phase is termed "cleavage" and results in splitting of that hexose into two 3 carbon molecules. One of those three carbon molecules is recycled back through the RuMP pathway and the other 3 carbon fragment is utilized for cell growth. In methanotrophs and methylotrophs the RuMP pathway may occur as one of three variants. However only two of these variants are commonly found. The FBP/TA (fructose bisphosphatase/Transaldolase) or the KDPG/TA (keto deoxy phosphogluconate/transaldolase) pathway. (Dijkhuizen L., G.E. Devries. The Physiology and biochemistry of aerobic methanol-utilizing gram negative and gram positive bacteria. In: Methane and Methanol Utilizers 1992, ed Colin Murrell and Howard Dalton Plenum Press NY).

The present strain is unique in the way it handles the "cleavage" steps where genes were found that carry out this conversion via fructose bisphosphate as a key intermediate. The genes for fructose bisphosphate aldolase and transaldolase were found clustered together on one piece of DNA. Secondly the genes for the other variant involving the keto deoxy phosphogluconate intermediate were also found clustered together. Available literature teaches that these organisms (methylotrophs and methanotrophs) rely solely on the KDPG pathway and that the FBP-dependent fixation pathway is utilized by facultative methylotrophs (Dijkhuizen et al., *supra*). Therefore the latter observation is expected whereas the former is not. The finding of the FBP genes in and obligate methane utilizing bacterium is both surprising and suggestive of utility. The FBP pathway is energetically favorable to the host microorganism due to the fact that less energy (ATP) is utilized than is utilized in the KDPG pathway. Thus organisms that utilize the FBP pathway may have an energetic advantage and growth advantage over those that utilize the KDPG pathway. This advantage may also be useful for energy-requiring production pathways in the strain. By using this pathway a methane-

utilizing bacterium may have an advantage over other methane utilizing organisms as production platforms for either single cell protein or for any other product derived from the flow of carbon through the RuMP pathway.

- Accordingly the present invention provides a method for the
- 5 production of a monoterpenes compound comprising providing a transformed C1 metabolizing host cell which
- (a) grows on a C1 carbon substrate selected from the group consisting of methane and methanol; and
  - (b) comprises a functional Embden-Meyerhoff carbon pathway,
- 10 said pathway comprising a gene encoding a pyrophosphate dependent phosphofructokinase enzyme; and
- (c) contains an endogenous source of geranyl diphosphate (GPP)

Isolation of C1 Metabolizing Microorganisms

- 15 The C1 metabolizing microorganisms of the present invention are ubiquitous and many have been isolated and characterized. A general scheme for isolation of these strains includes addition of an inoculum into a sealed liquid mineral salts media, containing either methane or methanol. Care must be made of the volume:gas ratio and cultures are typically
- 20 incubated between 25-55°C. Typically, a variety of different methylotrophic bacteria can be isolated from a first enrichment, if it is plated or streaked onto solid media when growth is first visible. Methods for the isolation of methanotrophs are common and well known in the art (See for example Thomas D. Brock in *Biotechnology: A Textbook of Industrial Microbiology*,
- 25 Second Edition (1989) Sinauer Associates, Inc., Sunderland, MA; Deshpande, Mukund V., *Appl. Biochem. Biotechnol.*, 36: 227 (1992); or Hanson, R.S. et al. *The Prokaryotes: a handbook on habitats, isolation, and identification of bacteria*; Springer-Verlag: Berlin, New York, 1981; Volume 2, Chapter 118).
- 30 As noted above, preferred C1 metabolizer is one that incorporates an active Embden-Meyerhoff pathway as indicated by the presence of a pyrophosphate dependent phosphofructokinase. It is contemplated that the present teaching will enable the general identification and isolation of similar strains. For example, the key characteristics of the present high
- 35 growth strain are that it is an obligate methanotroph, using only either methane or methanol as a sole carbon source and possesses a functional Embden-Meyerhoff, and particularly a gene encoding a pyrophosphate

dependent phosphofructokinase. Methods for the isolation of methanotrophs are common and well known in the art (See for example Thomas D. Brock *supra* or Deshpande, *supra*). Similarly, pyrophosphate dependent phosphofructokinase has been well characterized in

- 5 mammalian systems and assay methods have been well developed (see for example Schlisselfeld et al. *Clin. Biochem.* (1996), 29(1), 79-83; Clark et al., *J. Mol. Cell. Cardiol.* (1980), 12(10), 1053-64. The contemporary microbiologist will be able to use these techniques to identify the present high growth strain.

10 Genes involved in Monoterpene synthesis

Many C1 metabolizing strains possess the ability to produce geranyl diphosphate (GPP) which is the substrate for monoterpene synthases. Where a host cell is employed that makes GPP it will only be necessary to introduce the specific terpene synthase for the production of 15 a specific monoterpene.

Many cyclic terpene synthases are known in the art and any one will be suitable for expression in the hosts of the present invention.

Limonene synthase is the most well characterized having been isolated from a variety of organisms including *Perilla frutescens* (Genbank

- 20 Acc #AF317695), *Arabidopsis* (Genbank Acc # AB005235), *Perilla citriodora* (Genbank Acc # AF241790), *Schizonepeta tenuifolia* (Genbank Acc # AF233894), *Abies grandis* (Genbank Acc # AF139207), *Mentha longifolia* (Genbank Acc # AF175323) and *Mentha spicata* (Genbank Acc # L13459). Any one of the known genes encoding limonene synthase

- 25 may be used for expression in the present invention where genes isolated form and *Mentha spicata* are preferred.

Other cyclic terpene synthases are known. For example bornyl diphosphate synthase has been isolated from *Salvia officinalis* (Genbank Acc # AF051900); 1,8-cineole synthase has been isolated from *Salvia*

- 30 *officinalis* (Genbank Acc # AF051899); phellandrene synthase has been isolated from *Abies grandis* (Genbank Acc # AF139205); sabinene synthase has been isolated from *Salvia officinalis* (Genbank Acc # AF051901); and pinene synthase has been isolated from *Artemisia annua* (Genbank Acc # AF276072), and *Abies grandis* (Genbank Acc # U87909).

Accordingly, suitable synthases for monoterpene expression include but not limited to limonene synthase, pinene synthase, bornyl

synthase, phellandrene synthase, cineole synthase, and sabinene synthase.

It will be appreciated that where GPP is present in the host cell, expression of a specific terpene synthase will generate the corresponding monoterpenes. So for example, the expression of limonene synthase will generate limonene, the expression of pinene synthase will generate pinene, the expression of sabinene synthase will generate sabinene, the expression of phellandrene synthase will generate  $\beta$ -phellandrene and the expression of bornyl diphosphate synthase will generate borneol (Figure 3).

In some instances the specific C1 metabolizing host cell may be lacking some or all the elements of the pathway necessary for the production of geranyl diphosphate (GPP). Alternatively some of the elements of this pathway may be rate limiting and require overexpression for effective synthesis of GPP. In these situations it may be necessary to introduce some or all of the GPP synthetic pathway genes or "upper pathway isoprenoid genes" into the host, or to introduce additional copies of existing genes in the pathway to regulate or increase the production of certain rate limiting steps of the pathway. GPP is the end product of a biosynthetic pathway that begins with the condensation of Glyceraldehyde-3P and pyruvate and ends with the condensation of isopentenyl diphosphate (IPP) and dimethylallyl-diphosphate to form GPP (Figure 3).

Many steps in isoprenoid pathways are known. For example, the initial steps of the alternate pathway involve the condensation of 2 carbons from pyruvate with C1 aldehyde group, D-glyceraldehyde 3-Phosphate to yield 5-carbon compound (D-1-deoxyxylulose-5-phosphate) (Figure 3 and Figure 4). Lois *et al.* has reported a gene, *dxs*, that encodes D-1-deoxyxylulose-5-phosphate synthase (DXS) that catalyzes the synthesis of D-1-deoxyxylulose-5-phosphate in *E. coli* (*Proc. Natl. Acad. Sci. USA* 95: 2105-2110 (1998)).

Next, the isomerization and reduction of D-1-deoxyxylulose-5-phosphate yields 2-C-methyl-D-erythritol-4-phosphate. One of the enzymes involved in the isomerization and reduction process is D-1-deoxyxylulose-5-phosphate reductoisomerase (DXR). Takahashi *et al.* reported that *dxr* gene product catalyzes the formation of 2-C-methyl-D-

erythritol-4-phosphate in the alternate pathway in *E. coli* (*Proc. Natl. Acad. Sci. USA* 95: 9879-9884 (1998)).

Steps converting 2-C-methyl-D-erythritol-4-phosphate to isopentenyl monophosphate are not well characterized although some

- 5 steps are known. 2-C-methyl-D-erythritol-4-phosphate is then converted into 4-diphosphocytidyl-2C-methyl-D-erythritol in a CTP dependent reaction by the enzyme encoded by non-annotated gene *ygbP*. Rohdich *et al.* reported YgbP, a protein in *E. coli* that catalyzes the reaction mentioned above. Recently, *ygbP* gene was renamed as *ispD* as a part of  
10 *isp* gene cluster (SwissProt#Q46893) (*Proc. Natl. Acad. Sci. USA* 96:11758-11763 (1999)).

Then the 2 position hydroxy group of 4-diphosphocytidyl-2C-methyl-D-erythritol can be phosphorylated in an ATP dependent reaction by the enzyme encoded by *ychB* gene. Luttgen *et al.* has reported the presence  
15 of YchB protein in *E. coli* that phosphorylates 4-diphosphocytidyl-2C-methyl-D-erythritol resulting in 4-diphosphocytidyl-2C-methyl-D-erythritol 2-phosphate. Recently, *ychB* gene was renamed as *ispE* as a part of *isp* gene cluster (SwissProt #P24209) (Luttgen *et al.*, *Proc. Natl. Acad. Sci. USA* 97:1062-1067 (2000)).

- 20 Herz *et al.* reported that *ygbB* gene product in *E. coli* converts 4-diphosphocytidyl-2C-methyl-D-erythritol 2-phosphate to 2C-methyl-D-erythritol 2,4-cyclodiphosphate. 2C-methyl-D-erythritol 2,4-cyclodiphosphate can be further converted into carotenoids in the carotenoid biosynthesis pathway (*Proc. Natl. Acad. Sci. USA* 97:2486-2490 (2000)). Recently, *ygbB* gene was renamed as *ispF* as a  
25 part of *isp* gene cluster (SwissProt #P36663).

- 30 The reaction catalyzed by YgbP enzyme is carried out in CTP dependent manner. Thus CTP synthase plays an important role in the isoprenoid pathway. PyrG encoded by *pyrG* gene in *E. coli* was determined to encode CTP synthase (Weng *et al.*, *J. Biol. Chem.*, 261:5568-5574 (1986)).

- 35 Followed by the reactions not yet characterized, isopentenyl monophosphate is formed. Isopentenyl monophosphate is converted to isopentenyl diphosphate (IPP, C5) by isopentenyl monophosphate kinase encoded by *ipk* gene that is identical to the above mentioned *ychB* (*ispE*) gene (Lange and Croteau, *Proc. Natl. Acad. Sci. USA* 96:13714-13719 (1999)). Isopentenyl diphosphate (IPP) is isomerized to dimethylallyl-

pyrophosphate (DMAPP) by IPP:DMAPP isomerase (IPP isomerase, EC 5.3.3.2) or isopentenyl diphosphate isomerase (idi). Alternatively, recent evidence suggests that DMAPP can be formed separately at an earlier step of the mevalonate-independent pathway (Cunningham *et al.*, *J. Bac.*

5 182 No. 20: 5841-5848(2000)), and that the enzyme encoded by *lytB* plays an essential role for this alternate route of DMAPP formation.

DMAPP and IPP are condensed by geranyltranstransferase (*ispA*) gene (Ohto *et al.*, *Plant Mol. Biol.* 40 (2), 307-321 (1999) to produce the linear C-10 compound geranyl diphosphate (GPP) which is the substrate for

10 monoterpane synthases.

Accordingly, where it is necessary to regulate or install elements of the pathway needed for the synthesis of GPP in any C1 metabolizer, genes, known in the art, encoding the enzymes selected the group consisting of Dxs (1-deoxyxylulose-5-phosphate synthase), Dxr (1-deoxyxylulose-5-phosphate reductoisomerase), IspD (2C-methyl-D-erythritol cytidyltransferase), IspE (4-diphosphocytidyl-2-C-methylerythritol kinase), IspF, (2C-methyl-d-erythritol 2,4-cyclodiphosphate synthase), PyrG (CTP synthase), IspA (Geranyltransferase or farnesyl diphosphate synthase) and LytB may be used in the present C1 metabolizer host cell.

15 20 Construction of a Recombinant C1 Metabolizer for Monoterpene Production

Methods for introduction of genes encoding the appropriate cyclic terpene synthase into a suitable methylotrophic host are common.

Microbial expression systems and expression vectors containing

25 regulatory sequences suitable for expression of heterologous genes in methylotrophs are known. Any of these could be used to construct chimeric genes for expression of the any of the above mentioned cyclic terpene synthases. These chimeric genes could then be introduced into appropriate methylotrophic hosts via transformation to provide high level expression of the enzymes.

30 Vectors or cassettes useful for the transformation of suitable host cells are available. For example several classes of promoters may be used for the expression of genes encoding cyclic terpene synthases in methylotrophs and methanotrophs including, but not limited to:

35 endogenous promoters such as the deoxy-xylulose phosphate synthase, methanol dehydrogenase operon promoter (Springer *et al.* (1998) *FEMS Microbiol Lett* 160:119-124) the promoter for polyhydroxyalkanoic acid

- synthesis (Foellner et al. *Appl. Microbiol. Biotechnol.* (1993) 40:284-291), or promoters identified from native plasmid in methylotrophs (EP 296484). In addition to these native promoters non-native promoters may also be used, as for example the the promoter for lactose operon *P<sub>lac</sub>* (Toyama et al. *Microbiology* (1997) 143:595-602; EP 62971) or a hybrid promoter such as *P<sub>trc</sub>* (Brosius et al. (1984) *Gene* 27:161-172). Similarly promoters associated with antibiotic resistance e.g. kanamycin (Springer et al. (1998) *FEMS Microbiol Lett* 160:119-124; Ueda et al. *Appl. Environ. Microbiol.* (1991) 57:924-926) or tetracycline (US 4824786) are also suitable.
- Once the specific regulatory element is selected the promoter-gene cassette can be introduced into methylotrophs on a plasmid containing either a replicon (Brenner et al. *Antonie Van Leeuwenhoek* (1991) 60:43-48; Ueda et al. *Appl. Environ. Microbiol.* (1991) 57:924-926) for episomal expression or homologous regions for chromosomal integration (Naumov et al. *Mol. Genet. Mikrobiol. Virusol.* (1986) 11:44-48).
- Typically the vector or cassette contains sequences directing transcription and translation of the relevant gene, a selectable marker, and sequences allowing autonomous replication or chromosomal integration. Suitable vectors comprise a region 5' of the gene which harbors transcriptional initiation controls and a region 3' of the DNA fragment which controls transcriptional termination. It is most preferred when both control regions are derived from genes homologous to the transformed host cell, although it is to be understood that such control regions need not be derived from the genes native to the specific species chosen as a production host.
- Where accumulation of a specific monterpene is desired it may be necessary to reduce or eliminate the expression of certain genes in the target pathway or in competing pathways that may serve as competing sinks for energy or carbon. Methods of down-regulating genes for this purpose have been explored. Where sequence of the gene to be disrupted is known, one of the most effective methods gene down regulation is targeted gene disruption where foreign DNA is inserted into a structural gene so as to disrupt transcription. This can be effected by the creation of genetic cassettes comprising the DNA to be inserted (often a genetic marker) flanked by sequence having a high degree of homology to a portion of the gene to be disrupted. Introduction of the cassette into the host cell results in insertion of the foreign DNA into the structural gene via

the native DNA replication mechanisms of the cell. (See for example Hamilton et al. (1989) *J. Bacteriol.* 171:4617-4622, Balbas et al. (1993) *Gene* 136:211-213, Gueldener et al. (1996) *Nucleic Acids Res.* 24:2519-2524, and Smith et al. (1996) *Methods Mol. Cell. Biol.* 5:270-277.)

5        Antisense technology is another method of down regulating genes where the sequence of the target gene is known. To accomplish this, a nucleic acid segment from the desired gene is cloned and operably linked to a promoter such that the anti-sense strand of RNA will be transcribed. This construct is then introduced into the host cell and the antisense strand  
10 of RNA is produced. Antisense RNA inhibits gene expression by preventing the accumulation of mRNA which encodes the protein of interest. The person skilled in the art will know that special considerations are associated with the use of antisense technologies in order to reduce expression of particular genes. For example, the proper level of  
15 expression of antisense genes may require the use of different chimeric genes utilizing different regulatory elements known to the skilled artisan.

Although targeted gene disruption and antisense technology offer effective means of down regulating genes where the sequence is known, other less specific methodologies have been developed that are not  
20 sequence based. For example, cells may be exposed to a UV radiation and then screened for the desired phenotype. Mutagenesis with chemical agents is also effective for generating mutants and commonly used substances include chemicals that affect non-replicating DNA such as HNO<sub>2</sub> and NH<sub>2</sub>OH, as well as agents that affect replicating DNA such as  
25 acridine dyes, notable for causing frameshift mutations. Specific methods for creating mutants using radiation or chemical agents are well documented in the art. See for example Thomas D. Brock in  
Biotechnology: A Textbook of Industrial Microbiology, Second Edition  
(1989) Sinauer Associates, Inc., Sunderland, MA., or Deshpande, Mukund  
30 V., *Appl. Biochem. Biotechnol.*, 36, 227, (1992).

Another non-specific method of gene disruption is the use of transposable elements or transposons. Transposons are genetic elements that insert randomly in DNA but can be latter retrieved on the basis of sequence to determine where the insertion has occurred. Both  
35 *in vivo* and *in vitro* transposition methods are known. Both methods involve the use of a transposable element in combination with a transposase enzyme. When the transposable element or transposon, is contacted with

a nucleic acid fragment in the presence of the transposase, the transposable element will randomly insert into the nucleic acid fragment. The technique is useful for random mutagenesis and for gene isolation, since the disrupted gene may be identified on the basis of the sequence of

- 5 the transposable element. Kits for *in vitro* transposition are commercially available (see for example The Primer Island Transposition Kit, available from Perkin Elmer Applied Biosystems, Branchburg, NJ; based upon the yeast Ty1 element; The Genome Priming System, available from New England Biolabs, Beverly, MA; based upon the bacterial transposon Tn7; 10 and the EZ::TN Transposon Insertion Systems, available from Epicentre Technologies, Madison, WI, based upon the Tn5 bacterial transposable element.

In the context of the present invention the disruption of certain genes in the terpenoid pathway may enhance the accumulation of specific

- 15 monoterpenes however, the decision of which genes to disrupt would need to be determined on an empirical basis. Candidate genes may include one or more of the prenyltransferase genes which, as described earlier, which catalyze the successive condensation of isopentenyl diphosphate resulting in the formation of prenyl diphosphates of various chain lengths (multiples 20 of C-5 isoprene units). Other candidate genes for disruption would include any of those which encode proteins acting upon the terpenoid backbone prenyl diphosphates.

#### Industrial Production

- Where commercial production of the instant proteins are desired a 25 variety of culture methodologies may be applied. For example, large-scale production of a specific gene product, overexpressed from a recombinant microbial host may be produced by both batch or continuous culture methodologies.

- A classical batch culturing method is a closed system where the 30 composition of the media is set at the beginning of the culture and not subject to artificial alterations during the culturing process. Thus, at the beginning of the culturing process the media is inoculated with the desired organism or organisms and growth or metabolic activity is permitted to occur adding nothing to the system. Typically, however, a "batch" culture 35 is batch with respect to the addition of carbon source and attempts are often made at controlling factors such as pH and oxygen concentration. In batch systems the metabolite and biomass compositions of the system

change constantly up to the time the culture is terminated. Within batch cultures cells moderate through a static lag phase to a high growth log phase and finally to a stationary phase where growth rate is diminished or halted. If untreated, cells in the stationary phase will eventually die. Cells in log phase are often responsible for the bulk of production of end product or intermediate in some systems. Stationary or post-exponential phase production can be obtained in other systems.

5 A variation on the standard batch system is the Fed-Batch system. Fed-Batch culture processes are also suitable in the present invention and  
10 comprise a typical batch system with the exception that the substrate is added in increments as the culture progresses. Fed-Batch systems are useful when catabolite repression is apt to inhibit the metabolism of the cells and where it is desirable to have limited amounts of substrate in the media. Measurement of the actual substrate concentration in Fed-Batch  
15 systems is difficult and is therefore estimated on the basis of the changes of measurable factors such as pH, dissolved oxygen and the partial pressure of waste gases such as CO<sub>2</sub>. Batch and Fed-Batch culturing methods are common and well known in the art and examples may be found in Thomas D. Brock in Biotechnology: A Textbook of Industrial  
20 Microbiology, Second Edition (1989) Sinauer Associates, Inc., Sunderland, MA., or Deshpande, Mukund V., *Appl. Biochem. Biotechnol.*, 36, 227, (1992), herein incorporated by reference.

Commercial production of the instant proteins may also be accomplished with a continuous culture. Continuous cultures are an open  
25 system where a defined culture media is added continuously to a bioreactor and an equal amount of conditioned media is removed simultaneously for processing. Continuous cultures generally maintain the cells at a constant high liquid phase density where cells are primarily in log phase growth. Alternatively continuous culture may be practiced with  
30 immobilized cells where carbon and nutrients are continuously added, and valuable products, by-products or waste products are continuously removed from the cell mass. Cell immobilization may be performed using a wide range of solid supports composed of natural and/or synthetic materials.  
35 Continuous or semi-continuous culture allows for the modulation of one factor or any number of factors that affect cell growth or end product concentration. For example, one method will maintain a limiting nutrient

such as the carbon source or nitrogen level at a fixed rate and allow all other parameters to moderate. In other systems a number of factors affecting growth can be altered continuously while the cell concentration, measured by media turbidity, is kept constant. Continuous systems strive

- 5 to maintain steady state growth conditions and thus the cell loss due to media being drawn off must be balanced against the cell growth rate in the culture. Methods of modulating nutrients and growth factors for continuous culture processes as well as techniques for maximizing the rate of product formation are well known in the art of industrial
- 10 microbiology and a variety of methods are detailed by Brock, *supra*.

Fermentation media in the present invention must contain suitable carbon substrates. The suitable carbon substrate may be one-carbon substrates such as methane or methanol for which metabolic conversion into key biochemical intermediates has been demonstrated. For example,

- 15 methylotrophic yeast are known to utilize the carbon from methylamine to form trehalose or glycerol (Bellion et al., *Microb. Growth C1 Compd.*, [Int. Symp.], 7th (1993), 415-32. Editor(s): Murrell, J. Collin; Kelly, Don P. Publisher: Intercept, Andover, UK). Similarly, various species of *Candida* will metabolize alanine or oleic acid (Sulter et al., *Arch. Microbiol.* 20 153:485-489 (1990)). Hence it is contemplated that the source of carbon utilized in the present invention may encompass a wide variety of carbon containing substrates and will only be limited by the choice of organism.

#### Description of the Preferred Embodiments

- 25 *Methylomonas 16a* was isolated from a pond sediment using methane as sole source of carbon and energy. Among the colonies that were able to grow using methane as a sole source of carbon and energy, *Methylomonas 16a* strain was chosen for its rapid growth rate and pink pigmentation indicating inherent isoprenoid pathway for carotenoids.

The carbon flux pathways in *Methylomonas 16a* were analyzed by 30 gene expression profiling and the presence and activity of the Embden-Meyerhoff pathway, comprising the presence of a functional pyrophosphate-linked phosphofructokinase enzyme as confirmed.

A truncated limonene synthase gene lacking the first 57 amino acids of the protein from *Mentha spicata* was obtained from pR58 plasmid.

- 35 The truncated limonene synthase gene was cloned into the broad host vector pTJS75:dxS:dxR:Tn5Kn. The resulting plasmid pDH3 was transferred into *Methylomonas 16a* by triparental conjugal mating with

fresh overnight cultures of *E. coli* helper pRK2013 and *E. coli* donor DH10B/pDH3.-Vector pTJS75:dxS:dxR:Tn5Kn was similarly transferred into *Methylomonas*. Cloning methods and triparental conjugal mating are well known in the art. The presence of limonene synthase gene is verified 5 using PCR.

The transformed culture of *Methylomonas* 16a was grown in airtight bottles to prevent the loss of volatile limonene compound. The compound produced by transformed *Methylomonas* 16a was extracted and analyzed by gas chromatography. The compound was confirmed to be limonene .  
10 when compared to standard limonene. Approximately 0.5ppm of limonene was detected from transformed culture.

#### EXAMPLES

The present invention is further defined in the following Examples. It should be understood that these Examples, while indicating preferred 15 embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions.  
20

#### GENERAL METHODS

Standard recombinant DNA and molecular cloning techniques used in the Examples are well known in the art and are described by Sambrook, J., Fritsch, E. F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, (1989) 25 (Maniatis) and by T. J. Silhavy, M. L. Bennan, and L. W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) and by Ausubel, F. M. et al., Current Protocols in Molecular Biology, pub. by Greene Publishing Assoc. and Wiley-Interscience (1987).

Materials and methods suitable for the maintenance and growth of bacterial cultures are well known in the art. Techniques suitable for use in the following examples may be found as set out in Manual of Methods for General Bacteriology (Phillipp Gerhardt, R. G. E. Murray, Ralph N. Costilow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg and G. Briggs Phillips, eds), American Society for Microbiology, Washington, DC. (1994)) or by Thomas D. Brock in Biotechnology: A Textbook of Industrial 35 Microbiology, Second Edition, Sinauer Associates, Inc., Sunderland, MA

(1989). All reagents, restriction enzymes and materials used for the growth and maintenance of bacterial cells were obtained from Aldrich Chemicals (Milwaukee, WI), DIFCO Laboratories (Detroit, MI), GIBCO/BRL (Gaithersburg, MD), or Sigma Chemical Company (St. Louis, MO) unless otherwise specified.

The meaning of abbreviations is as follows: "h" means hour(s), "min" means minute(s), "sec" means second(s), "d" means day(s), "mL" means milliliters, "L" means liters.

Microbial Cultivation and Preparation of Cell Suspensions, and associated analyses.

*Methylomonas* 16a is typically grown in serum stoppered Wheaton bottles using a gas/liquid ratio of at least 8:1 (i.e. 20 mL of Nitrate liquid media) media in a Wheaton bottle (Wheaton Scientific, Wheaton IL) of 160 mL total volume. The standard gas phase for cultivation contained 25% methane in air. These conditions comprise growth conditions and the cells are referred to as growing cells. In all cases the cultures were grown at 30°C with constant shaking in a Lab-Line rotary shaker unless otherwise specified.

Cells obtained for experimental purposes were allowed to grow to maximum optical density (O.D. 660 ~ 1.0). Harvested cells were obtained by centrifugation in a Sorval RC-5B centrifuge using a SS-34 rotor at 6000 rpm for 20 min. These cell pellets were resuspended in 50 mM HEPES buffer pH 7. These cell suspensions are referred to as washed, resting cells.

Microbial growth was assessed in all experiments by measuring the optical density of the culture at 660 nm in an Ultrospec 2000 UV/Vis spectrophotometer (Pharmacia Biotech, Cambridge England) using a 1 cm light path cuvet. Alternatively microbial growth was assessed by harvesting cells from the culture medium by centrifugation as described above and resuspending the cells in distilled water with a second centrifugation to remove medium salts. The washed cells were then dried at 105°C overnight in a drying oven for dry weight determination.

Methane concentration was determined as described by Emptage et al. (1997 *Env. Sci. Technol.* 31:732-734), hereby incorporated by reference.

Nitrate medium for *Methylomonas 16A*

Nitrate liquid medium, also referred to herein as "defined medium" was comprised of various salts mixed with solution 1 as indicated below or where specified the nitrate was replaced with 15 mM ammonium chloride.

- 5      Solution 1 Composition for 100 fold concentrated stock solution of trace minerals.

	<i>MW</i>	<i>Conc.</i> ( <i>mM</i> )	<i>g per L</i>
Nitriloacetic acid	191.1	66.9	12.8
CuCl <sub>2</sub> x 2H <sub>2</sub> O	170.48	0.15	0.0254
FeCl <sub>2</sub> x 4H <sub>2</sub> O	198.81	1.5	0.3
MnCl <sub>2</sub> x 4H <sub>2</sub> O	197.91	0.5	0.1
CoCl <sub>2</sub> x 6H <sub>2</sub> O	237.9	1.31	0.312
ZnCl <sub>2</sub>	136.29	0.73	0.1
H <sub>3</sub> BO <sub>3</sub>	61.83	0.16	0.01
Na <sub>2</sub> MoO <sub>4</sub> x 2H <sub>2</sub> O	241.95	0.04	0.01
NiCl <sub>2</sub> x 6H <sub>2</sub> O	237.7	0.77	0.184

- 10     Mix the gram amounts designated above in 900 mL of H<sub>2</sub>O, adjust to pH=7, and add H<sub>2</sub>O to an end volume of 1 L. Keep refrigerated.

Nitrate liquid medium:

	<i>MW</i>	<i>Conc.</i> ( <i>mM</i> )	<i>g per L</i>
NaNO <sub>3</sub>	84.99	10	0.85
KH <sub>2</sub> PO <sub>4</sub>	136.09	3.67	0.5
Na <sub>2</sub> SO <sub>4</sub>	142.04	3.52	0.5
MgCl <sub>2</sub> x 6H <sub>2</sub> O	203.3	0.98	0.2
CaCl <sub>2</sub> x 2H <sub>2</sub> O	147.02	0.68	0.1
1 M HEPES (pH 7) Solution 1	238.3		50 mL 10 mL

- 15     Dissolve in 900 mL H<sub>2</sub>O. Adjust to pH=7, and add H<sub>2</sub>O to give 1 L.  
For agar plates: Add 15 g of agarose in 1 L of medium, autoclave, let cool down to 50°C, mix, and pour plates.

Nitrate and Nitrite Assays

1 mL samples of cell culture were taken and filtered through a 0.2 micron Acrodisc filter to remove cells. The filtrate from this step contains the nitrite or nitrate to be analyzed. The analysis was performed on a Dionex ion chromatograph 500 system (Dionex, Sunnyvale, CA) with an AS3500 autosampler. The column used was a 4 mm Ion-Pac AS11-HC separation column with an AG-AC guard column and an ATC trap column. All columns are provided by Dionex.

- 5                   The mobile phase was a potassium hydroxide gradient from 0 to 10 50 mM potassium hydroxide over a 12 min time interval. Cell temperature was 35°C with a flow rate of 1 mL/min.

Microarray of gene expressionAmplification of DNA regions for the construction of DNA

microarray: Specific primer pairs were used to amplify each protein specifying ORF of *Methylomonas* sp. strain 16a. Genomic DNA (10-30 ng) was used as the template. The PCR reactions were performed in the presence of HotStart Taq™ DNA polymerase (Qiagen, Valencia, CA) and the dNTPs (Gibco BRL Life Science Technologies, Gaithersburg, MD). Thirty-five cycles of denaturation at 95°C for 30 sec, annealing at 20 55°C for 30 sec and polymerization at 72°C for 2 min were conducted. The quality of PCR reactions was checked with electrophoresis in a 1% agarose gel. The DNA samples were purified by the high-throughput PCR purification kit from Qiagen.

- 25                  Arraying amplified ORFs. Before arraying, an equal volume of DMSO (10 µL) and DNA (10 µL) sample was mixed in 384-well microtiter plates. A generation II DNA spotter (Molecular Dynamics, Sunnyvale, CA) was used to array the samples onto coated glass slides (Telechem, Sunnyvale, CA). Each PCR product was arrayed in duplicate on each slide. After cross-linking by UV light, the slides were stored under vacuum 30 in a desiccator at room temperature.

- 35                  RNA isolation: *Methylomonas* 16a was cultured in a defined medium with ammonium or nitrate (10 mM) as nitrogen source under 25% methane in air. Samples of the minimal medium culture were harvested when the O.D. reaches 0.3 at A<sub>600</sub> (exponential phase). Cell cultures were harvested quickly and ruptured in RLT buffer [Qiagen RNeasy Mini Kit, Valencia, CA] with a beads-beater (Bio101, Vista, CA). Debris was pelleted by centrifugation for 3 min at 14,000 x g at 4°C. RNA isolation was completed using the protocol supplied with this kit. After on-column

DNAase treatment, the RNA product was eluted with 50-100 µL RNAase-free. RNA preparations were stored frozen at either -20 or -80°C.

Synthesis of fluorescent cDNA from total RNA. RNA samples (7 to 15 µg) and random hexamer primers (6 µg; Gibco BRL Life Science

- 5 Technologies) were diluted with RNAase-free water to a volume of 25 µL. The sample was denatured at 70°C for 10 min and then chilled on ice for 30 seconds. After adding 14 µL of labeling mixture, the annealing was accomplished by incubation at room temperature for 10 min. The labeling mixture contained 8 µL of 5x enzyme buffer, 4 µL DTT (0.1M), and 2 µL of  
10 20x dye mixture. The dye mixture consisted of 2 mM of each dATP, dGTP, and dTTP, 1 mM dCTP, and 1 mM of Cy3-dCTP or Cy5-dCTP. After adding 1 to 1.5 µL of SuperScript II reverse transcriptase  
15 (200 units/mL, Life Technologies Inc., Gaithersburg, MD), cDNA synthesis was allowed to proceed at 42°C for 2 hr. The RNA was removed by adding 2 µL NaOH (2.5 N) to the reaction. After 10 min of incubation at 37°C, the pH was adjusted with 10 µL of HEPES (2M). The labeled cDNA was then purified with a PCR purification kit (Qiagen, Valencia, CA). Labeling efficiency was monitored using either A<sub>550</sub> for Cy3 incorporation, or A<sub>650</sub> for Cy5.

20 Fluorescent labeling of genomic DNA. Genomic DNA was nebulized to approximately 2 kb pair fragments. Genomic DNA (0.5 to 1 µg) was mixed with 6 µg of random hexamers primers (Gibco BRL Life Science Technologies) in 15 µL of water. The mix was denatured by put at boiling water for 5 minutes. Then anneal on ice for 30 sec before put at  
25 room temperature. Then 2 µL 5x Buffer 2 (Gibco BRL) and 2 ul dye mixture were added. The component of dye mixture and the labeling procedure are the same as described above for RNA labeling, except that the Klenow fragment of DNA polymerase I (5 µg/µL, Gibco BRL Life Science Technologies) was used as the enzyme. After incubation 37 °C  
30 for 2 hr, the labeled DNA probe was purified using a PCR purification kit (Qiagen, Valencia, CA).

35 Hybridization and washing. Slides were first incubated with prehybridization solution containing 3.5xSSC (BRL, Life Technologies Inc., Gaithersburg, MD), 0.1% SDS (BRL, Life Technologies Inc., Gaithersburg, MD), 1% bovine serum albumin (BSA, Fraction V, Sigma, St. Louis, MO). After prehybridization, hybridization solutions (Molecular Dynamics) containing labeled probes was added to slides and covered with cover

slips. Slides were placed in a humidified chamber in a 42°C incubator. After overnight hybridization, slides were initially washed for 5 min at room temperature with a washing solution containing 1xSSC, 0.1% SDS and 0.1xSSC, 0.1% SDS. Slides were then washed at 65°C for 10 min with 5 the same solution for three times. After washing, the slides were dried with a stream of nitrogen gas.

Data Collection and Analysis. The signal generated from each slide was quantified with a laser scanner (Molecular Dynamics, Sunnyvale, CA). The images were analyzed with ArrayVision 4.0 software (Imaging 10 Research, Inc., Ontario, Canada). The raw fluorescent intensity for each spot was adjusted by subtracting the background. These readings were exported to a spreadsheet for further analysis.

#### EXAMPLE 1

##### ISOLATION OF METHYLOMONAS 16A

15 The original environmental sample containing the isolate was obtained from pond sediment. The pond sediment was inoculated directly into defined medium with ammonium as nitrogen source under 25% methane in air. Methane was the sole source of carbon and energy. Growth was followed until the optical density at 660 nm was stable 20 whereupon the culture was transferred to fresh medium such that a 1:100 dilution was achieved. After 3 successive transfers with methane as sole carbon and energy source the culture was plated onto growth agar with ammonium as nitrogen source and incubated under 25% methane in air. Many methanotrophic bacterial species were isolated in this manner. 25 However, *Methylomonas 16a* was selected as the organism to study due to the rapid growth of colonies, large colony size, ability to grow on minimal media, and pink pigmentation indicative of an active biosynthetic pathway for carotenoids.

#### EXAMPLE 2

##### RAPID GROWTH ON METHANE IN MINIMAL MEDIUM

30 *Methylomonas 16a* grows on the defined medium comprised of only minimal salts, a culture headspace comprised of methane in air. Methane concentrations for growth but typically are 5-50% by volume of the culture headspace. No organic additions such as yeast extract or vitamins are 35 required to achieve growth shown in Figure 5. Figure 5 shows the growth of 16a compared to the growth of *Methylococcus capsulatus* under identical growth conditions. i.e. minimal medium with 25% methane in air

as substrate. The data indicates *Methylomonas* 16a doubles every 2-2.5 h whereas *Methylococcus capsulatus* doubles every 3.5 h with methane as substrate. With methanol as substrate doubling times on methanol are 2.5-3 for *Methylomonas* 16a and 4.5-5 for *Methylococcus capsulatus*. Cell 5 densities are also significantly higher for *Methylomonas* 16a growing on methane. *Methylococcus capsulatus* is a widely utilized methanotroph for experimental and commercial purposes.

### EXAMPLE 3

#### COMPARISON OF GENE EXPRESSION LEVELS IN THE ENTNER DOUDEROFF PATHWAY AS COMPARED WITH THE EMBDEN MEYEROF PATHWAY

Example 3 presents microarray evidence for the use of the Embden-Meyerhoff pathway in the 16a strain.

Figure 6 shows the relative levels of expression of genes for the 15 Entner-Douderoff pathway and the Embden-Meyerhof pathway. The relative transcriptional activity of each gene was estimated with DNA microarray as described previously (Wei, et al., 2001. *Journal of Bacteriology*. 183:545-556).

Specifically, a single DNA microarray containing 4000 ORFs (open 20 reading frames) of *Methylomonas* sp. strain 16a was hybridized with probes generated from genomic DNA and total RNA. The genomic DNA of 16a was labeled with Klenow fragment of DNA polymerase and fluorescent dye Cy-5, while the total RNA was labeled with reverse transcriptase and Cy-3. After hybridization, the signal intensities of both Cy-3 and Cy-5 for 25 each spot in the array were quantified. The intensity ratio of Cy-3 and Cy-5 was then used to calculate the fraction of each transcript (in percentage) with the following formula: (gene ratio/sum of all ratio) x 100. The value obtained reflects the relative abundance of mRNA of an individual gene. Accordingly, transcriptional activity of all the genes represented by the 30 array can be ranked based on its relative mRNA abundance in a descending order. For example, mRNA abundance for the methane monooxygenase was ranked #1 because its genes had the highest transcriptional activity when the organism was grown with methane as the carbon source (Figure 6).

The genes considered "diagnostic" for Entner-Douderoff are the 35 6 phosphogluconate dehydratase and the 2 keto-3-deoxy-6-phosphogluconate aldolase. Phosphofructokinase and fructose

bisphosphate aldolase are "diagnostic" of the Embden-Meyerhof sequence. Numbers in Figure 6 next to each step indicate the relative expression level of that enzyme. For example the most highly expressed enzyme in the cell is the methane monooxygenase (ranked #1). The next most highly

- 5 expressed is the methanol dehydrogenase (ranked #2). Messenger RNA transcripts of Phosphofructokinase (ranked #232) and fructose  
bisphosphate aldolase (ranked #65) were in higher abundance than those  
for glucose 6 phosphate dehydrogenase (ranked #717), 6  
phosphogluconate dehydratase (ranked #763) or the 2-keto-3-deoxy-6-  
10 gluconate aldolase. The data suggests that the Embden-Meyerhof  
pathway enzymes are more strongly expressed than the Entner-Douderoff  
pathway enzymes. This result is surprising and counter to existing beliefs  
on the central metabolism of methanotrophic bacteria (Reference book  
pages in. The physiology and biochemistry of aerobic methanol-utilizing  
15 gram-negative and gram- positive bacteria In: Methane and Methanol  
Utilizers, Biotechnology Handbooks 5. 1992. Eds: Colin Murrell, Howard  
Dalton. Pp. 149-157.

EXAMPLE 4

DIRECT ENZYMATIC EVIDENCE FOR A PYROPHOSPHATE-LINKED  
PHOSPHOFRUCTOKINASE

20 Example 4 shows the evidence for the presence of a pyrophosphate-linked phosphofructokinase enzyme in the current strain which would confirm the functionality of the Embden-Meyerhof pathway in the present strain.

25 Phosphofructokinase activity was shown to be present in *Methylomonas 16a* by using the coupled enzyme assay described below. Assay conditions are given in Table 2 below. This assay was further used to assay the activity in a number of other Methanotrophic bacteria as shown below in Table 3. The data in Table 3 show known ATCC strains  
30 tested for phosphofructokinase activity with ATP or pyrophosphate as phosphoryl donor. These organisms were classified as either Type I or Type X ribulose monophosphate-utilizing strains or Type II serine utilized.  
Coupled Assay Reactions

35 Phosphofructokinase reaction was measured by a coupled enzyme assay. Phosphofructokinase reaction was coupled with fructose 1,6, biphosphate aldolase followed by triosephosphate isomerase. The enzyme activity was measured by the disappearance of NADH.

Specifically, the enzyme phosphofructokinase catalyzes the key reaction converting Fructose 6 phosphate and pyrophosphate to Fructose 1,6 bisphosphate and orthophosphate. Fructose-1,6-bisphosphate is cleaved to 3-phosphoglyceraldehyde and dihydroxyacetonephosphate by fructose 1,6-bisphosphate aldolase. Dihydroxyacetonephosphate is isomerized to 3-phosphoglyceraldehyde by triosephosphate isomerase. Glycerol phosphate dehydrogenase plus NADH and 3-phosphoglyceraldehyde yields the alcohol glycerol-3-phosphate and NAD. Disappearance of NADH is monitored at 340 nm using spectrophotometer (UltraSpec 4000, Pharmacia Biotech).

Table 2  
Assay Protocol

Reagent	Stock solution (mM)	Volume ( $\mu$ l) per 1mL total reaction volume	Final assay concentration (mM)
Tris-HCl pH 7.5	1000	100	100
MgCl <sub>2</sub> · 2 H <sub>2</sub> O	100	35	3.5
Na <sub>4</sub> P <sub>2</sub> O <sub>7</sub> · 10H <sub>2</sub> O or ATP	100	20	2
Fructose-6-phosphate	100	20	2
NADH	50	6	0.3
Fructose bisphosphate aldolase	100 ( units/mL)	20	2 (units)
Triose phosphate isomerase/glyceraldehyde phosphate dehydrogenase	(7.2 units/ $\mu$ l) (0.5 units/ $\mu$ l)	3.69	27 units 1.8 units
KCl	1000	50	50
H <sub>2</sub> O		adjust to 1mL	
Crude extract		0-50	

Table 3  
Comparison Of Pyrophosphate Linked And ATP Linked  
Phosphofructokinase Activity In Different Methanotrophic Bacteria

Strain	Type	Assimilation Pathway	ATP-PFK umol NADH/min/mg	Ppi-PFK umol NADH/min/mg
Methylomonas 16a ATCC PTA 2402	I	Ribulose monophosphate	0	2.8
Methylomonas agile ATCC 35068	I	Ribulose monophosphate	0.01	3.5
Methylobacter Whittenbury ATCC 51738	I	Ribulose monophosphate	0.01	0.025
Methylomonas clara ATCC 31226	I	Ribulose monophosphate	0	0.3
Methylococcus capsulatus ATCC 19069	X	Ribulose monophosphate	0.02	3.6
Methylosinus sporium ATCC 35069	II	Serine	0.07	0.4

5

Several conclusions may be drawn from the data presented above. First, it is clear that ATP (which is the typical phosphoryl donor for phosphofructokinase) is essentially ineffective in the phosphofructokinase reaction in methanotrophic bacteria. Only inorganic pyrophosphate was found to support the reaction in all methanotrophs tested. Secondly not all methanotrophs contain this activity. The activity was essentially absent in *Methylobacter whittenbury* and in *Methylococcus capsulatus*. Intermediate levels of activity were found in *Methylomonas clara* and *Methylosinus sporium*. These data show that many methanotrophic bacteria may contain a hitherto unreported phosphofructokinase activity. It may be inferred from this that methanotrophs containing this activity have an active Embden-Meyerhof pathway.

#### EXAMPLE 5

#### GROWTH YIELD AND CARBON CONVERSION BY METHYLOMONAS

20

#### 16A

Growth yield and carbon conversion efficiency were compared for *Methylomonas 16a* and *Methylococcus capsulatus*. These strains were

chosen because 16a contains high levels of phosphofructokinase and *M. capsulatus* is essentially devoid of the enzyme activity. It was contemplated that if *Methylomonas* 16a could utilize the more energetically favorable Embden-Meyerhof pathway and *Methylococcus capsulatus* could only use the Entner-Douderoff pathway the superior energetics of the present *Methylomonas* 16a strain would be reflected in cellular yields and carbon conversion efficiency. This difference in energetic efficiency would only be apparent under energy-limiting conditions. These conditions were achieved in this experiment by limiting the amount of oxygen in each culture to only 10% (vol/vol) instead of 20%. Under these oxygen limiting conditions the strain that produces the most energy from aerobic respiration on methane will produce more cell mass.

Cells were grown as 200 mL cultures 500 mL serum-stoppered Wheaton bottles. The headspace in the bottles was adjusted to 25% methane and 10% oxygen. The defined medium formulation is the same in both cases.

Table 4

Yield Of *Methylomonas* 16a Cells Versus *Methylococcus Capsulatus* Cells  
Under Oxygen Limitation.

Strain	$Y_{CH_4}$ g dry wt/mol	G dry wt/g $CH_4$	Carbon Conversion Efficiency (CCE)%
<i>Methylomonas</i> 16a	16.7 +/- 0.5	1.04	64
<i>Methylococcus</i> <i>capsulatus</i>	10.3 +/- 0.3	0.64	40

Yield determination: Yield was measured by growing triplicate cultures in 500 mL bottles on defined medium with ammonium as nitrogen source under oxygen limitation. This was done by using 300 mL of culture with a 300 mL headspace of 25% methane and 10% oxygen the balance being nitrogen. At the end of growth (i.e. stationary phase) residual methane in the headspace was determined by gas chromatography. The cells were collected by centrifugation washed with distilled water and dried overnight in a drying oven before being weighed.

Carbon conversion efficiency is a measure of how much carbon is assimilated into cell mass. It is calculated assuming a biomass composition of CH<sub>2</sub>O<sub>0.5</sub>N<sub>0.25</sub>:

*Methylomonas 16a* : 16 g/mol methane X (1 g dry wt/g methane) /

5 25 g/ mol biomass

*M. capsulatus* 16 g/mol methane X (0.64 g dry wt/g methane) /

25 g / mol biomass

These data (in Table 4) show that *Methylomonas 16a* produced significantly more cell mass than did the *Methylococcus capsulatus* strain

10 under growth conditions that were identical except for the temperature. *Methylococcus capsulatus* grows optimally at 45°C whereas *Methylomonas* is grown at 33°C. It may be inferred from the data that the presence of the more energy-yielding Embden-Meyerhof pathway confers a growth advantage to *Methylomonas 16a*.

15 Table 5 presents the theoretical calculations showing ATP yield as a function of carbon assimilation pathway with the carbon output being normalized to pyruvate in all cases (The physiology and biochemistry of aerobic methanol-utilizing gram-negative and gram- positive bacteria In: Methane and Methanol Utilizers, Biotechnology Handbooks 5. 1992. Eds: Colin Murrell, Howard Dalton. Pp. 149-157). Table 5 shows the amount of ATP that is produced or consumed for every three molecules of carbon (as formaldehyde or carbon dioxide) for serine cycle, xylulose monophosphate cycle and ribulose monophosphate cycle pathways. The latter pathway, as discussed is typically thought to exist as the 2-keto-3deoxy-6-phosphogluconate /transaldolase (KDPGA/TA) variant. These data shows that in fact the fructose bisphosphate aldolase/transaldolase (FBPA/TA) variant is likely to exist in the methanotrophs. The energetic repercussion of this is the net production of an additional 1 ATP for methanotrophs if they possess an ATP linked phosphofructokinase and an additional 2 ATPs for the pyrophosphate-linked enzyme. It is therefore expected that *Methylomonas 16a* derives and additional 2 ATP per 3 carbons assimilated and that this may explain the greater yield and carbon efficiency of the strain versus *Methylococcus capsulatus*.

Table 5  
Energetics of Methanotrophic bacteria utilizing different carbon  
assimilation mechanisms

Organism	Cycle	C1 unit fixed	Product	Variant	ATP	NADPH
Bacteria	RuMP	3CH <sub>2</sub> O	Pyruvate	FBPA/TA	+1	+1
Methylomonas	RuMP /Serine	3CH <sub>2</sub> O	Pyruvate	FBPA/TA	+1(+2*)	+1
Bacteria	RuMP	3CH <sub>2</sub> O	Pyruvate	KDPGA/TA	0	+1
Methylococcus	RuMP/RuBP	3CH <sub>2</sub> O	Pyruvate	KDPGA/TA	0	+1

5 \* Based on PPI dependent phosphofructokinase

EXAMPLE 6  
Plasmid Construction.

The plasmid pR58 contains *Mentha spicata* limonene synthase  
10 gene which carries a deletion of the first 57 amino acids of the enzyme  
(Williams *et al*, *Biochemistry* 1998, 37, 12213-12220). pR58 was digested  
with restriction enzymes *NdeI* and *BamH1*, releasing the truncated  
limonene synthase. The 1638 base pair gene was purified by agarose gel  
electrophoresis and QIAEX II Gel Extraction (Qiagen Cat.# 20021). The  
15 4 base pair 5' overhang (5'-TATG-3') gap was filled using T4 DNA  
polymerase (Gibco BRL Cat.# 18005-017), maintaining the methionene  
start codon immediately preceding the arginine codon which corresponds  
to amino acid number 58 in the native gene. Concurrently, the 3' *BamH1*  
overhang was filled in by the same enzyme without interruption of the  
20 native limonene synthase stop codon. The blunt ended insert was purified  
using the QIAquick PCR Purification Kit (Qiagen Cat# 28104).

The vector pTJS75:dxS:dxR:Tn5Kn is a derivative of RK2, a broad-host-range plasmid (*J. Bact.*, 164, 446-455) modified to include kanamycin resistance and two *Methylomonas* genes of interest: *dxs*, encoding  
25 5'deoxy-D-xylulose synthase and *dxr*, encoding 5'deoxy-D-xylulose reductoisomerase. This vector was prepared for ligation by digesting with *XbaI*, blunting the overhangs with T4 DNA Polymerase, dephosphorylating with Calf Intestinal Alkaline Phosphatase (GibcoBRL), and purifying on a QIAquick PCR Purification column.

30 The limonene synthase gene insert was ligated into the vector and electroporated into *E. coli* electroMAX DH10B cells (Gibco BRL). The

resulting plasmid, designated pDH3, contains the limonene synthase gene flanked on the 5'end by *dxS* and on the 3'end by *dxR* (Figure 1).

Plasmid pDH3 was transferred into *Methylomonas sp. 16a* by triparental conjugal mating. Fresh overnight cultures of *E. coli* helper

- 5 pRK2013 and *E. coli* donor DH10B/pDH3 along with vector (pTJS75:*dxS*:*dxR*:Tn5Kn) control grown in LB (Luria-Bertani medium) with kanamycin (50 µg/mL) were washed three times in plain LB, and resuspended in a volume of LB representing approximately a 60-fold concentration of the original culture volume. Recipient cells, a rifampicin
- 10 resistant subculture of *Methylomonas sp. 16a* (designated 16aA), were grown for 48 hours in BTZ-3 (Table 1) with rifampicin (50µg/mL) under 25% methane, washed three times in BTZ-3, and resuspended in a volume of BTZ-3 representing approximately a 150-fold concentration of the original culture volume. The resulting donor, helper, and recipient cell
- 15 pastes were combined on the surface of BTZ-3 agar plates containing 0.5% (w/v) yeast extract in ratios of 1:1:2 respectively. Plates were maintained at 25% methane for 16-24 hours to allow conjugation to occur. Cell pastes were harvested and resuspended in BTZ-3. Serial dilutions were plated on BTZ-3 agar with rifampicin (50 µg/mL) and kanamycin
- 20 (50 µg/mL). Resulting colonies were patched to new selective plates. Colonies which grew the second time were then transferred to liquid BTZ-3 with rifampicin (50 µg/mL) and kanamycin (50 µg/mL). The presence of the limonene synthase gene in the 16a conjugates was verified by PCR using primer set 5'-atgagacgatccggaaactacaaccc-3' (SEQID NO:1) and 5'-
- 25 tcatgcaaaggctcgaaaaataagggtctgg-3' (SEQID NO:2) which anneals to the N- and C-terminus of the limonene synthase respectively. The primer set 5'-atgattgaacaagatggattgc-3'(SEQID NO:3) and 5'-aagcttcaaaagaactcgtc-3' (SEQID NO:4) was used to detect kanamycin resistance gene as a control.

30

#### EXAMPLE 7

##### Limonene Detection.

Liquid cultures of *Methylomonas sp. 16a* transconjugants were grown in airtight bottles for approximately 48 hours in BTZ-3 supplemented with rifampicin (50 µg/mL), kanamycin (50 µg/mL) and 25% methane.

35

Contents of the culture vessel were harvested for limonene and analyzed by gas chromatography (GC) analysis. Briefly, ethyl acetate was injected through the septa in order to extract limonene from both the headspace

and from the liquid. Samples were agitated for 5 min., transferred to appropriate centrifuge tubes and centrifuged to achieve complete separation of the organic and aqueous phases. The organic phase was removed and loaded onto a DB-1 column (30M/0.25 ID/0.25 film thickness, 5 J&W Scientific, Folsom, CA).

The initial oven temperature was set at 50°C and temperature was increased to 250°C at a rate of 10 degrees per minute. After 5 minutes at 250°C, the samples were injected into the column. A Septum purge was applied for one minute at a rate of 2.4 mL per minute. The runs were

- 10 carried out at a head pressure of 8 psi and flow rate of 24.5 mL per minute. Limonene used as a standard was purchased from Sigma-Aldrich (Cat#18,316-4). Samples analyzed by GC/MS are identical to those analyzed by GC alone. GC/MS instrument type: ProSpec. Source type: Combined-EI/CI SRC. Electron energy: 69.0 eV. Multiplier voltages: #1: 15 198.6 V, #2: 250.0 V, #3: 250.0 V, #4: 250.0 V. Emission current: 0.0 mA Trap current: 444.1 µA. Ion repeller: 5.7 V. Source temp: 225.3°C. Extraction heater current: 0.0 mA. Slit criteria: Source: 47.8%, Collector: 50.4%, Alpha: 100%, Z2 Restrict: 0.0%, Z3 Restrict: 0.0%, Z4 Restrict: 0.0%. Lens Criteria: Ion Energy: 1.4, Focus#1: 2.2 V, Beam Centre: -48.7 20 V, Focus #2: 5.8 V, Y-Focus: 4.4 V, Y-Def#1: -88.0 V, Z-Def#1: 7.5 V, Rotate #1: 0.0 V, Z-Def#2: -12.6 V, Z-Focus #2: 6.0V, Rotate #2: -13.6 V, Curve #2: -4.0V, Curve #3: -8.6 V, Rotate #3: -9.0 V, Z-Focus #3: 9.3 V, Z-Def #3: 0.0 V, Rotate #4: 0.0 V. Magnet parameters: IMR 1160.16, Standard coil. No ramped parameters. HP6890 GC parameters:
- 25 Automatic restart, Capillary line temp (1): 240.0°C, Capillary line temp (2): 240.0°C, Reentrant temp: 240.0°C, Maximum oven temp: 325.2°C, Equilibrium time: 0.2 min, Oven Temperature Ramps: Temp #1: 50.0°C, Time #1: 1.0 min, Rate #1: 10.0°C/min, Temp#2: 200.0°C, Time #2: 1.0 min, Rate #2: 10.0°C/min, Temp #3: 250.0°C, Time #3: 10.0 min, Rate #3: 30 0.0°C/min. Injectors: "A": Active, Injector "A" temp: 270.0°C, Injection type: splitless, Purge "A" on time: 1.0 min, Purge "A" Flow: 1.0 mL/min, Col 1 Const Flow: 1.0 mL/min. "B": Active, Injection type: Cool-on-Col, "B" Injector Temp Ramps: Temp #1: 50.0°C, Time #1: 650.0 min, Rate #1: 0.0°C/min, Col 1 Const flow: 2.0mL/min.
- 35 *Methylomonas sp. 16a* which received the limonene synthase gene produced limonene (~0.5 ppm). In the GC analysis, chromatograms of these extracts display a peak which is superimposable with that of the

limonene standard (Figure 2). As expected, *Methylomonas sp. 16a* alone or *Methylomonas sp. 16a* containing vector pTJS75:dxS:dxR:Tn5Kn did not produce any detectable limonene. The presence or absence of limonene was further confirmed by GC/MS. The limonene peak from GC 5 generated identical fragmentation patterns as the limonene standard with a signature peak at 68 and predicted MW at 136. The level of production was comparable in the two assays.

CLAIMS

What is claimed is:

1. A method for the production of a monoterpene comprising:
  - a) providing a transformed C1 metabolizing host cell comprising:
    - (i) suitable levels of geranyl pyrophosphate; and
    - (ii) at least one isolated nucleic acid molecule encoding a cyclic terpene synthase under the control of suitable regulatory sequences;
  - b) contacting the host cell of step (a) under suitable growth conditions with an effective amount of a C1 carbon substrate whereby a monoterpene compound is produced.
2. A method according to Claim 1 wherein the C1 carbon substrate is selected from the group consisting of methane, methanol, formaldehyde, formic acid, methylated amines, methylated thiols, and carbon dioxide.
3. A method according to Claim 1 wherein the C1 metabolizing host cell is a methylotroph selected from the group consisting of *Methylomonas*, *Methylobacter*, *Methylcoccus*, *Methylosinus*, *Methylocystis*, *Methylomicrobium*, *Methanomonas*, *Methylophilus*, *Methylobacillus*, *Methylobacterium*, *Hypomicrobium*, *Xanthobacter*, *Bacillus*, *Paracoccus*, *Nocardia*, *Arthrobacter*, *Rhodopseudomonas*, *Pseudomonas*, *Candida*, *Hansenula*, *Pichia*, *Torulopsis*, and *Rhodotorula*.
4. A method according to Claim 1 wherein C1 metabolizing host is a methanotroph.
5. A method according to Claim 4 wherein the methanotroph is selected from the group consisting of *Methylomonas*, *Methylobacter*, *Methylcoccus*, *Methylosinus*, *Methylocystis*, *Methylomicrobium*, and *Methanomonas*.
6. A method according to Claim 2 wherein the C1 carbon substrate is selected from the group consisting of methane and methanol and the C1 metabolizing host cell is a methanotroph selected from the group consisting of *Methylomonas*, *Methylobacter*, *Methylcoccus*, *Methylosinus*, *Methylocystis*, *Methylomicrobium*, and *Methanomonas*.
7. A method according to Claim 6 wherein the obligate methanotroph is a high growth methanotrophic strain which comprises a functional Embden-Meyerof carbon pathway, said pathway comprising a gene encoding a pyrophosphate dependent phosphofructokinase enzyme.

8. A method according to Claim 7 wherein the high growth methanotrophic bacterial strain optionally contains at least one gene encoding a fructose bisphosphate aldolase enzyme.
9. A method according to Claim 7 wherein the high growth methanotrophic bacterial strain optionally contains a functional Entner-Douderoff carbon pathway.
- 5 10. A method according to Claim 9 wherein the high growth methanotrophic bacterial strain optionally contains at least one gene encoding a keto-deoxy phosphogluconate aldolase.
- 10 11. A method according to Claim 10 wherein the high growth methanotrophic bacterial strain is *methylomonas* 16a having the ATCC designation ATCC PTA 2402.
- 15 12. A method according to Claim 1 wherein the cyclic terpene synthase is selected from the group consisting of limonene synthase, pinene synthase, bornyl synthase, phellandrene synthase, cineole synthase, and sabinene synthase.
13. A method according to Claim 1 wherein the monoterpene is selected from the group consisting of limonene, pinene, bornyl diphosphate,  $\beta$ -phellandrene, 1,8-cineole, and sabinene.
- 20 14. A method according to Claim 1 wherein the cyclic terpene synthase is limonene synthase, the monoterpene is limonene and the recombinant host is *Methylomonas*.
15. A method according to Claim 14 wherein the limonene synthase has the amino sequence as set forth in SEQ ID NO:6.
- 25 16. A method according to Claim 14 wherein the limonene synthase is encoded by the gene as described in the sequences selected from the group consisting of Genbank Acc #AF317695, Genbank Acc # AB005235, Genbank Acc # AF241790 , Genbank Acc # AF233894 , Genbank Acc # AF139207, Genbank Acc # AF175323 and Genbank Acc # L13459.
- 30 17. A method according to Claim 13 wherein the pinene synthase is encoded by the gene as described in sequences selected from the group consisting of Genbank Acc # AF276072 ,and Genbank Acc # U87909.
18. A method according to Claim 13 wherein the bornyl synthase is encoded by the gene as described in Genbank Acc # AF051900.
- 35 19. A method according to Claim 13 wherein the phellandrene synthase is encoded by the gene as described in Genbank Acc # AF139205

20. A method according to Claim 13 wherein the cineole synthase is encoded by the gene as described in Genbank Acc # AF051899.

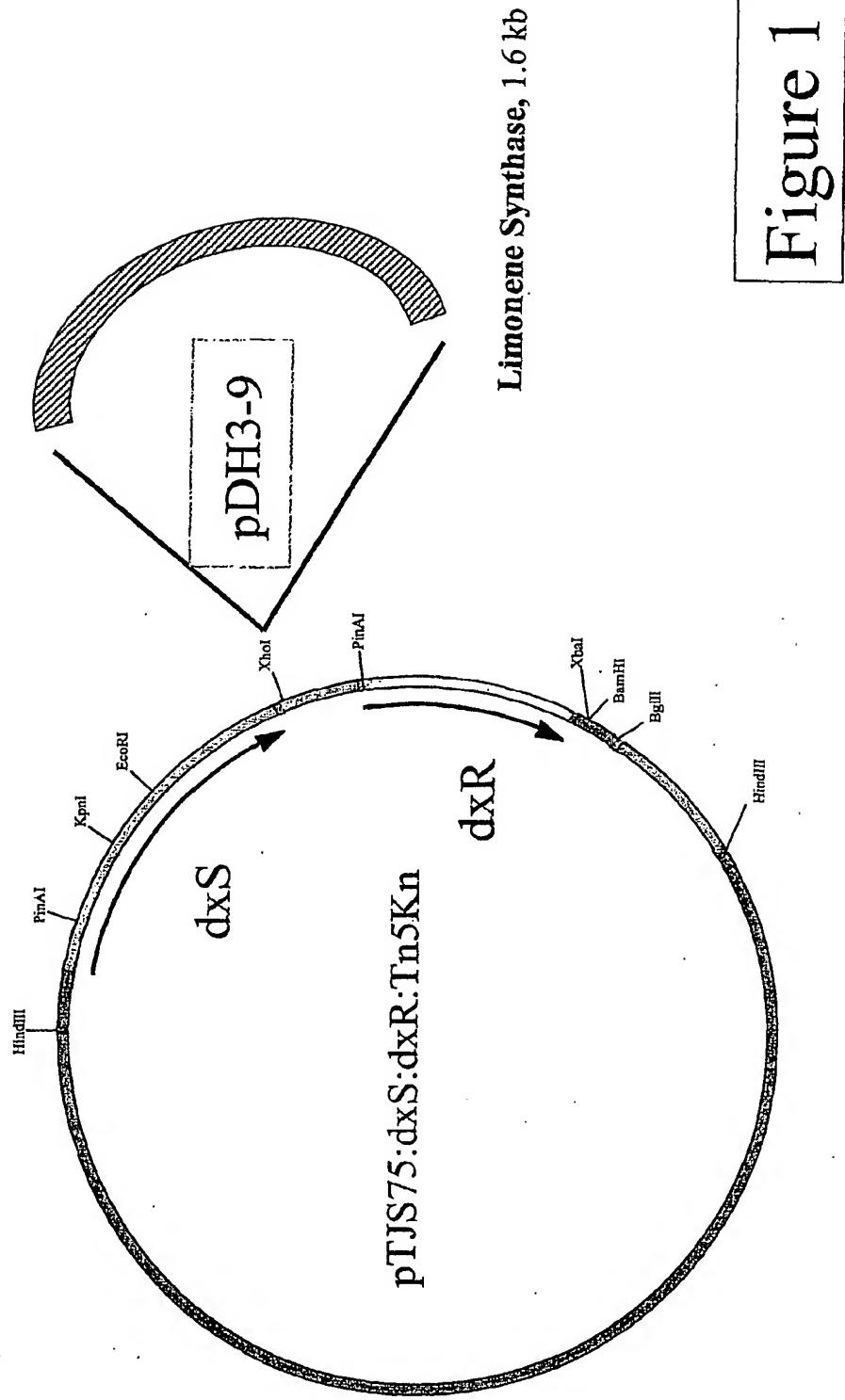
21. A method according to Claim 13 wherein the sabinene synthase is encoded by the gene as described in Genbank Acc

5 # AF051901

22. A method according to Claim 1 wherein the suitable levels of geranyl pyrophosphate are provided by the expression heterologus upper pathway isoprenoid pathway genes.

23. A method according to Claim 22 wherein said upper pathway  
10 isoprenoiod genes encod enzymes selected from the group consisting of D-1-deoxyxylulose-5-phosphate synthase (DXS); D-1-deoxyxylulose-5-phosphate reductoisomerase (DXR); 2C-methyl-d-erythritol cytidylyltransferase (IspD), 4-diphosphocytidyl-2-C-methylerythritol kinase (IspE), 2C-methyl-d-erythritol 2,4-cyclodiphosphate synthase (IspF), CTP  
15 synthase (IspA) and Geranyltranstransferase (PyrG).

# Expression vector and construct



## Gas Chromatography: Limonene Production in *Methylomonas*

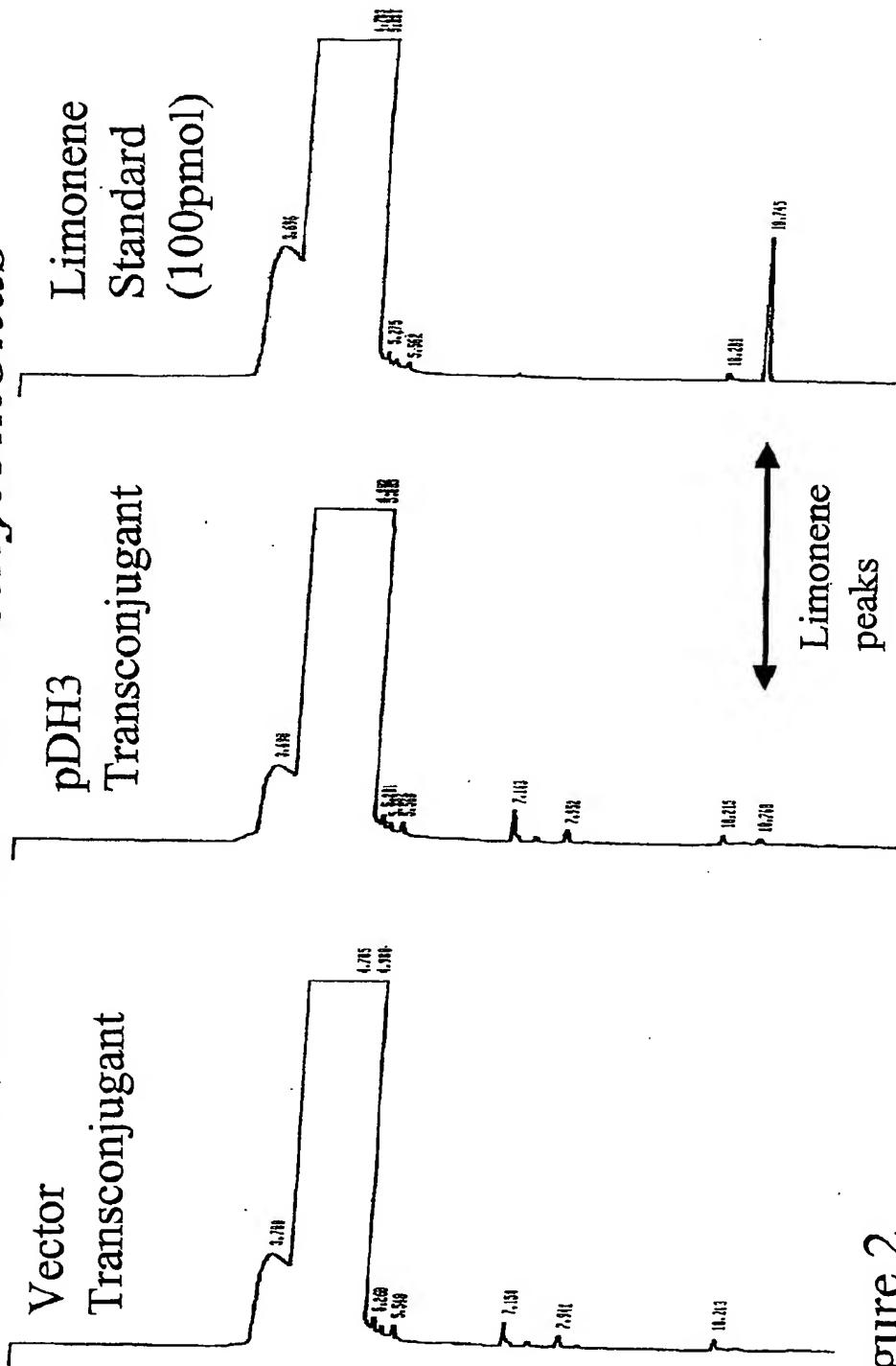


Figure 2

Figure 3

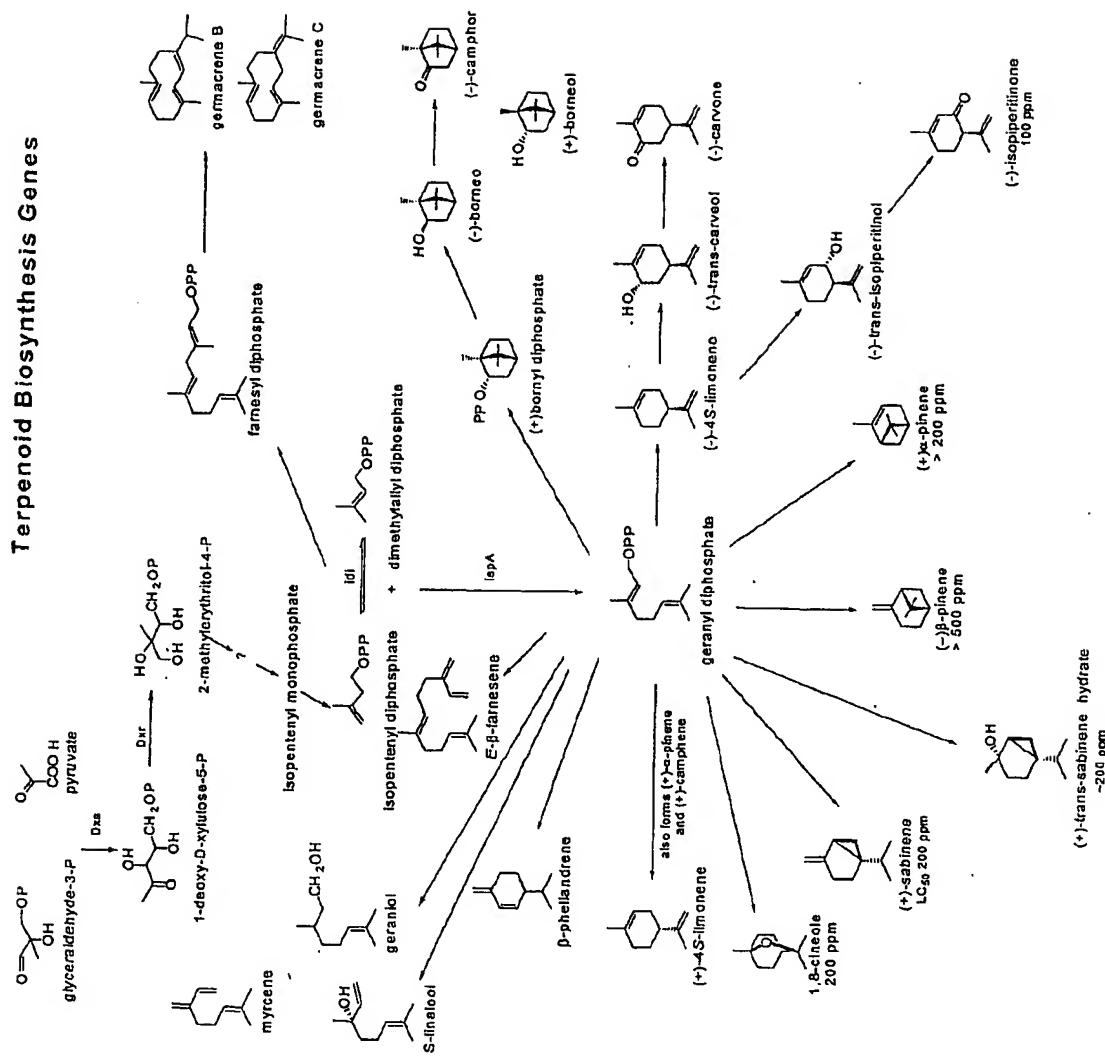


FIGURE 4

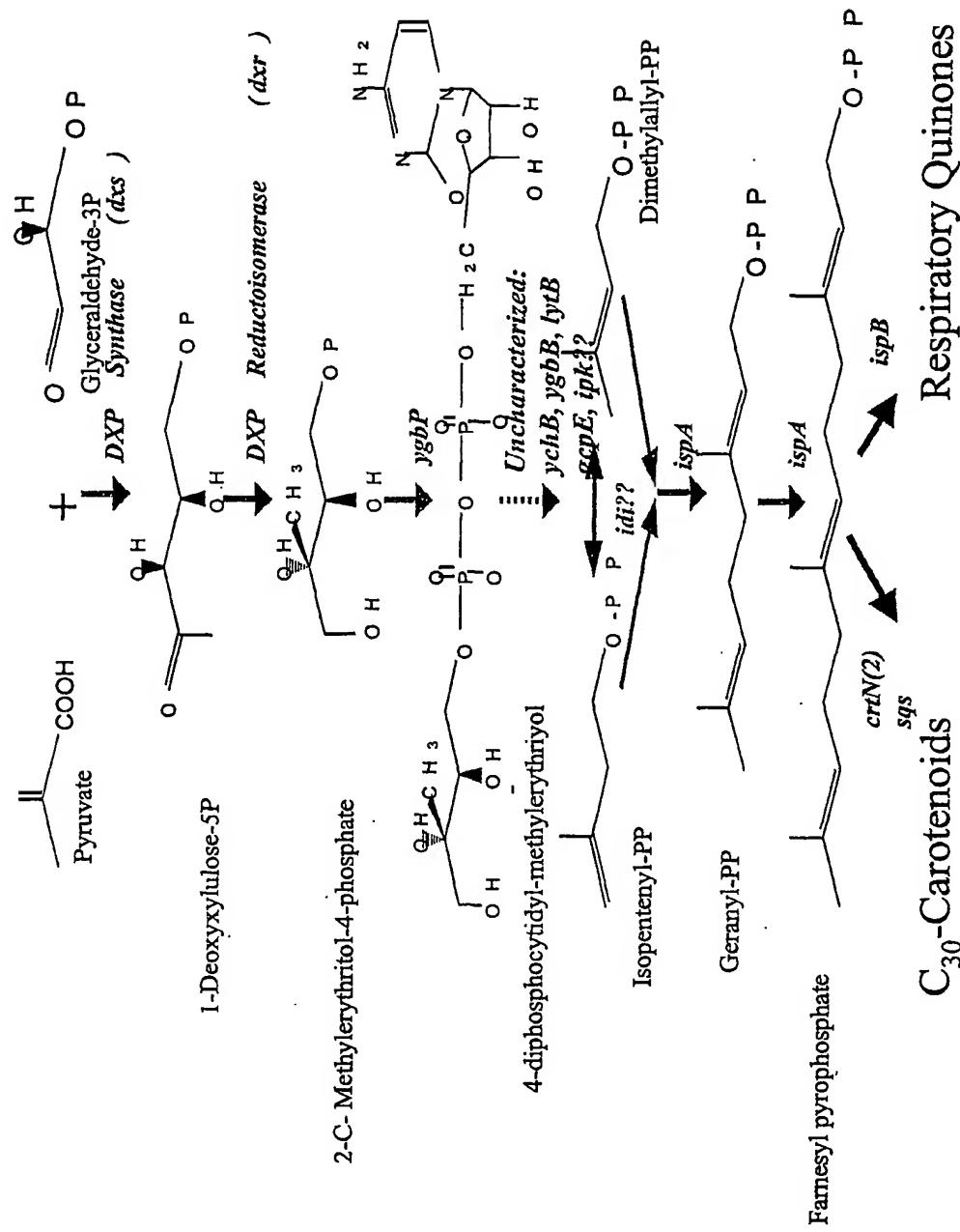


Figure 5

16a vs. MCapsulatus Growth Curve

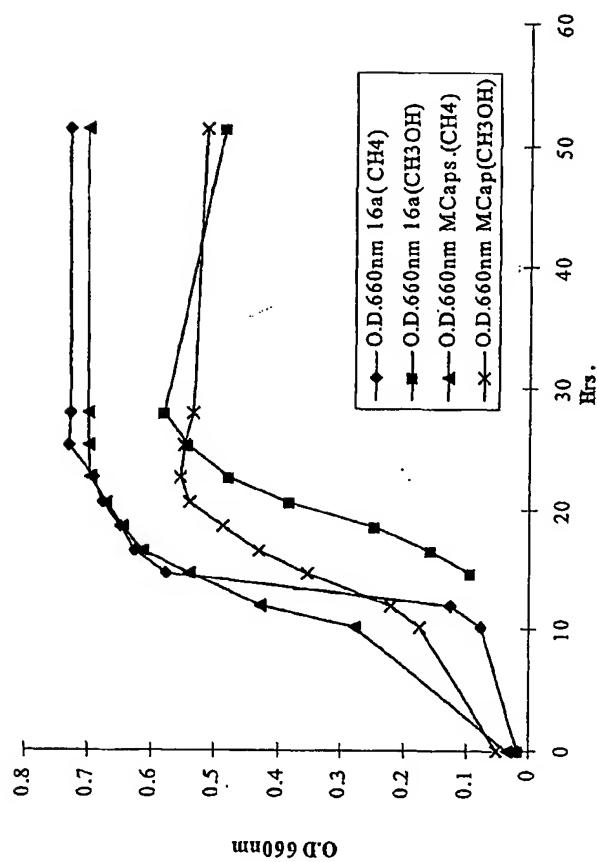
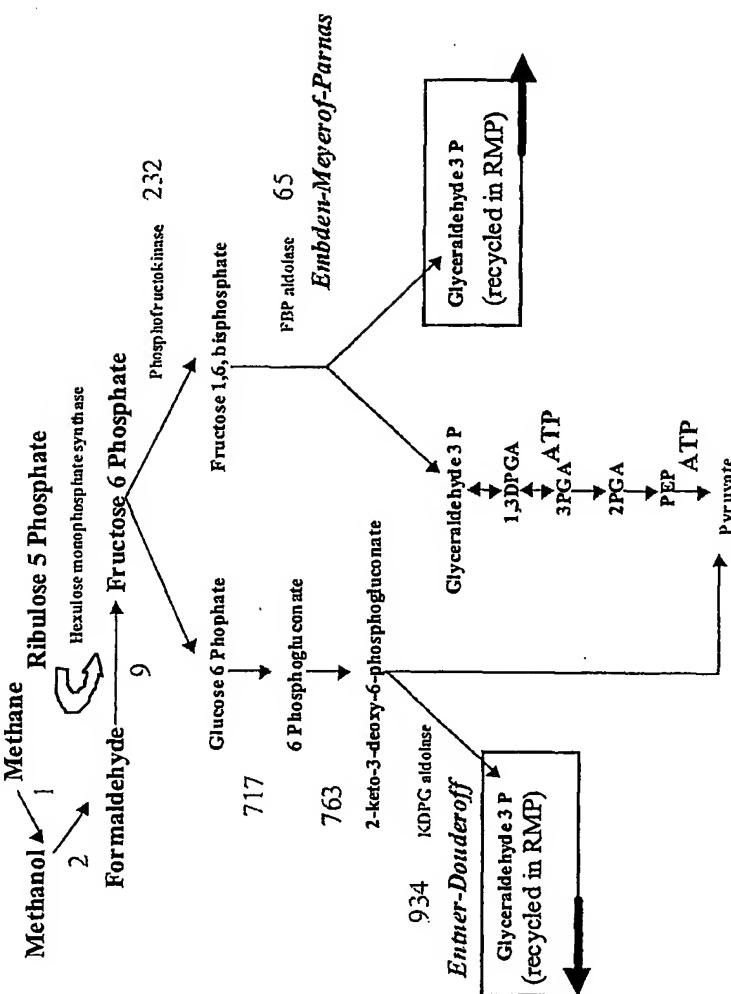


FIGURE 6



## SEQUENCE LISTING

V

<110> E.I. du Pont de Nemours and Company

<120> Production of Monoterpene

<130> CL1809

<150> 60/229,907

<151> 2000-09-0

<150> 60/229,858

<151> 2000-09-01

<160> 7

<170> Microsoft Office 97

<210> 1

<211> 26

<212> DNA

<213> Primer

<400> 1

atgagacgat ccggaaaacta caaccc

26

<210> 2

<211> 29

<212> DNA

<213> Primer

<400> 2

tcatgcaaag ggctcgaata aggttctgg

29

<210> 3

<211> 22

<212> DNA

<213> Primer

<400> 3

atgattgaac aagatggatt gc

22

<210> 4

<211> 21

<212> DNA

<213> Primer

<400> 4

aagctttcaa aagaactcgt c

21

<210> 5

<211> 11575

<212> DNA

<213> Plasmid

<400> 5

tcccgtggcg tcgaaagtgc ggcaccatag gtatcagtca ccgcgatgag atcccttacc 60

attccagagt ctggcggttg attattaatt tgctgatata gagcctcago ccgctggcga 120

aattcattac gtaaatcaaa ggcttcaggt cgggtaatt taaaactaag ctgaatgatt 180

ttctggagat agcggctgcc atcttcgata ttcaagcgcat gttcaacggc atgagtgata 240

atctgcctgt cataacagag aatatggta aagcggggca gatcggtcac tgcacgcaca 300

agcctgaaca cttccgcccc ac ctggatggc tccagtcggc ccagatcatc catgacaaca 360

atgaacttca gatccagact caccagttgt cctgcaattt cagcccgaag cttgcgtgt 420

ttcgtactcg gctggtttga aaccctgtgcg cgctcaagg atgatcccgg catcttgc 480

ttcttcttgc cgttactggc ggccttcggc ggcatgatgt tgctggcaca ctccatgtc 540

ggcttcgaag ccaaaaaccgc gttcttgatc caggtcggcc ataccttgcattt gggcgatattc 600

tcgctgatcc tggcctgcgg tcgctggctg gaactcaagc tcgattctcc cggcaaaaat 660

attgcgggtt ttatccagt gttcgccctg tttcaaatacg gcgtcatcct gatgttctac	720
cgtgaaccct tgtactgatt atgaaactga ccaccgacta tcccttgctt aaaaacatcc	780
acacgcggc ggacatacgc ggcgtgtcca aggaccagct ccagcaactg gctgacgagg	840
tgcgcggcta tctgacccac acggtcagca tttccggcgg ccattttgcg gccggcctcg	900
gcaccgtgga actgaccgtg gccttgatt atgtgttcaa taccggcgtc gatcagttgg	960
tctggacgt gggccatcag gcctatccgc acaagattct gaccggcgc aaggagcga	1020
tgccgaccat tcgcacccctg ggcggggtgt cagccttcc ggccggggac gagagcgaat	1080
acgatgcctt cggcgctggc cattccagca cctcgatcag cgccggactg ggcattggca	1140
ttgcgtcgca gtcgcggc gaagacaaga agatggtagc catcatcgcc gacggttcca	1200
tcaccggcgg catggcctat gaggcgatga atcatgccgg cgatgtgaat gccaacctgc	1260
tggtgatctt gaacgacaac gatatgtcga tctcgccgccc gtcggggcgt atgaacaatt	1320
atctgaccaa ggtgttgtcg agcaagttt attcgctggt gcgggaagag agcaagaaag	1380
ctctggccaa gatgcccgtcg gtgtggAAC tggcgcccaa gaccgaggaa cacgtgaagg	1440
gcatgatcgt gccccgtacc ttgttcgagg aattggcctt caattatttc ggcggatcg	1500
acggccatga tgcgtcgatg ctgggtgtca ccctggaaaa tctgaaggat ttgaccggc	1560
cggtatccct gcatgtggtg accaagaagg gcaaaaggcta tgccgcagcc gagaaagacc	1620
cgttggccta ccattggcggtcg ccggctttcg atccgaccaa ggatttcctg cccaaaggcgg	1680
cgccgtcgcc gcatccgacc tataccgagg tggtcgcccg ctggctgtgc gacatggcgg	1740
ctcaagacga ggcgttgtcg ggcattcacgc cggcgatgcg cgaaggctct ggtttgtgg	1800
aattctcaca gaaatttccg aatcgctatt tcgatgtcgc catcgccgag cagcatgcgg	1860
tgaccttggc cgccggccag gcctgccagg ggcggccagg ggtgggtggcg atttattcca	1920
ccttcctgca acgcggttac gatcgttgc tccacgacgt ggccttgcag aacttagata	1980
tgctctttgc actggatcgt gcccggcttgg tggccggcagg tggaccgacc catgctggcg	2040
cctttgatta cagctacatg cgctgtattc cgaacatgc gatcatggct ccagccgacg	2100
agaacgagtg caggcagatg ctgaccaccg gcttccaaca ccatggcccg gttcgggtgc	2160
gctatcccgcg cggcaaaggc cccggggcgg caatcgatcc gaccctgacc ggcgtggaga	2220
tcggcaaggc cgaagtcaaga caccacggca ggcgcacgc cattctggcc tggggcagca	2280
tggtcacgcc tgccgtcgaa gcccggcaagc agctggcgc gacgggtggtg aacatgcgtt	2340
tcgtcaagcc gttcgatcaa gccttggtgc tggaaatttgc caggacgcac gatgtttcg	2400
tcaccgtcgaa ggaaaacgtc atcgccggcg ggcgtggcag tgccatcaac accttcctgc	2460

aggcgacaaa	ggtgctgatg	ccggtctgca	acatoggcct	gcccgaccgc	ttcgctgagc	2520
aaggtagtcg	cgaggaattt	ctcagcctgg	tcggcctcga	cagcaagggc	atcctcgcca	2580
ccatcgaaaca	gttttgcgtt	taaacttgcc	gatgctggaa	atcattcaac	tgccagtcc	2640
gaacgacaac	tcgaggacat	cagtgcattat	ttcgctggca	aaaaatgggg	caaggacaaa	2700
ctcgccctg	aaatcagccc	tggcaaaacc	gtgcaaggca	tgtatggtgc	attggcttca	2760
gcgatgattt	gcfgatagg	tttgcgcgtt	tattacggct	tttcggcctt	ggaatcggt	2820
ggcgccgaat	tggcggtcct	gatgtcgata	gatttgcgtt	tttgcgttgt	gttgaccgtg	2880
ctggtatcca	tttacggcga	tttgcgttgtt	agtctggtca	agcgaatcaa	aggcgtaag	2940
gatagtggca	ccttgcgttgc	gggtcatggc	ggtatccctcg	atagggtgga	cagcatcatt	3000
gcggcggcac	cgttttctta	tgccggtata	gtgctgatcg	gacggagcgt	attcgaatga	3060
aaggattttt	catattgggc	gctaccgggt	cgatcggtgt	cagcacgctg	gatgtcgttt	3120
ccaggcatcc	ggataaatat	caagtcgttgc	cgctgaccgc	caacggcaat	atcgacgcatt	3180
tgtatgaaca	atgcctggcc	caccatccgg	agtatgcgggt	ggtggtcatg	gaaagcaagg	3240
tagcagagtt	caaacagcgc	attgcgcctt	cgccggtagc	ggatatcaag	gtcttgcgg	3300
gtagcgaggc	cttgcacacag	gtggccacgc	tggaaaacgt	cgatacggtg	atggcgctta	3360
tcgtcggcgc	ggccggattt	ttgcccacct	tggccgcggc	caaggccggc	aaaaccgtgc	3420
tgttggccaa	caaggaagcc	ttggtgatgt	cgggacaaat	cttcatgcag	gccgtcagcg	3480
attccggcgc	tgttttgctg	ccgatagaca	gcfagcacaa	cgccatcttt	cagtgcattgc	3540
cggcggtta	tacgcccaggc	catacagcca	aacaggcgcg	ccgcatttttta	ttgaccgctt	3600
ccgggtggccc	atttcgacgg	acgcccata	aaacgttgtc	cagcgtcacg	ccggatcagg	3660
ccgttgcacca	tcctaaatgg	gacatggggc	gcaagatttc	ggtcgattcc	gccaccatga	3720
tgaacaaagg	tctcgaaactg	atcgaaaccc	gcttgcgtt	caacatggag	cccgaccaga	3780
ttgaagtctgt	cattcatccg	cagagcatca	ttcattcgat	ggtggactat	gtcgatggtt	3840
cggttttggc	gcagatgggt	aatcccgaca	tgcgcacgcc	gatagcgcac	gcgtatggct	3900
ggccggaaacg	ctttgactct	ggtgtggcgc	cgctggatat	tttgcgttgt	gggcacatgg	3960
atttcgaaaa	acccgacttg	aaacggtttc	cttgcgttgt	attggcttac	gaagccatca	4020
agtctggtgg	aattatgcac	acggatttgc	acgcagccaa	tgaatttgc	gtcgaaacgt	4080
ttttaaatga	agaagtcaaa	ttcactgaca	tcgcggcata	catcgagcgc	agcatggccc	4140
agtttaaacc	ggacgatgcc	ggcagcctcg	aattggttt	gcaggccgat	caagatgcgc	4200
gcgagggtggc	tagagacatc	atcaagaccc	tggtagctt	atggaaaccc	ttcacaccct	4260

gttttattcc atcggtgcga tcgcatattct gggtgcctct agatcggttc cgtcgacact 4320  
 gcagagcttgc cagtggcatt acatggcgat agcttagactg ggccggttta tggacagcaa 4380  
 gcaaccggaa attgccagct ggggcgcctt ctggtaagggt tgggaagccc tgcaaaatgaa 4440  
 actggatggc tttcttgcccg ccaaggatct gatggcgcaag gggatcaaga tctgatcaag 4500  
 agacaggatg aggatcgttt cgcatgattt aacaagatgg attgcacgca gggtctccgg 4560  
 ccgcgttgggt ggagaggcta ttccggctatg actgggcaca acagacaatc gggtgtctg 4620  
 atgcgcgcgt gtccggctg tcagcgcaagg ggccggccgt tcttttgc aagaccgacc 4680  
 tgtccgggtgc cctgaatgaa ctgcaggacg aggcaagcg gctatcggtt ctggccacga 4740  
 cgggcgttcc ttgcgcagct gtgcgtcgact ttgtcaactga agcgggaagg gactggctgc 4800  
 tattggcga agtgcggggg caggatctcc tgtcatctca ccttgcctt gcccggaaaag 4860  
 tatccatcat ggctgatgca atgcggcgcc tcgatacgct tgatccggct acctgccccat 4920  
 tcgaccacca agcggaaacat cgcatcgagc gagcacgtac tcggatggaa gcccggcttg 4980  
 tcgatcagga tgatctggac gaagagcatc aggggcgtc gccagccgaa ctgtcgcca 5040  
 ggctcaaggc gcgcgtgcgc gacggcgagg atctcgctgt gacccatggc gatgcctgct 5100  
 tgccgaatat catggtgaa aatggccgtt tttctggatt catcgactgt ggccggctgg 5160  
 gtgtggcgga ccgctatcag gacatagcgt tggctaccgg tgatattgct gaagagctt 5220  
 gcggcgaatg ggctgaccgc ttccctcggtc tttacggat cccgcgtccc gattcgacgc 5280  
 gcatcgccctt ctatcgccctt cttgacgagt tcttctgaaa gcttgcgtc cattttggg 5340  
 gtgaggccgt tcgcggccga ggggcgcagc ccctgggggg atgggaggcc cgcgttagcg 5400  
 ggccggggagg gttcgagaag gggggcacc ccccttcggc gtgcgcggc acgcgcacag 5460  
 ggccgcagccc tggtaaaaaa caaggtttat aaatatttgtt taaaaagcag gttaaaagac 5520  
 aggttagcgg tggccgaaaa acggggcgga aacccttgca aatgtggat tttctgcctg 5580  
 tggacagccc ctcaaatgtc aatagggtcg cccctcatct gtcagcactc tgcccctcaa 5640  
 gtgtcaaggaa tcgcgcctt catctgtcag tagtcgcgc cctcaagtgtt caataccgca 5700  
 gggcacttat ccccaggctt gtccacatca tctgtggaa actcgctaa aatcaggcgt 5760  
 tttcgccgat ttgcgaggctt ggccagctcc acgtcgccgg ccgaaatcga gcctgcctt 5820  
 catctgtcaa cggccgcgcg ggtgagtcgg cccctcaagt gtcaacgtcc gcccctcatc 5880  
 tgcgtgttgc gccaagttt tccgcgaggat tccacaacg ccggccggccg cgggtgtctcg 5940  
 cacacggctt cgacggcggtt tctggcgctt ttgcaggggcc atagacggcc gccagccag 6000  
 cggcgaggc aaccagccccg gtgagcgtcg gaaagggtcg acggatctt tccgctgcat 6060

aaccctgctt	cgggtcatt	atagcgattt	tttcggata	tccatccctt	ttcgacat	6120
atacaggatt	ttgc当地	gttcgttag	acttcctt	gttatccaa	cggcgtcagc	6180
cgggcaggat	aggtgaagta	ggcccacccg	cgagcgggtg	ttc当地	actgtccctt	6240
attcgacact	ggc当地	aacggaaatc	ctgctctg	aggctggccg	gctaccgccc	6300
gc当地	tgagggcaag	cggatggctg	atgaaaccaa	gcc当地	aagggcagcc	6360
cacctatcaa	ggtgtactgc	ttccagac	aacgaagagc	gattgaggaa	aaggcggcgg	6420
cggccggcat	gagc当地	gc当地	tggccgtcgg	ccagggtac	aaaatcacgg	6480
gc当地	ctatgagcac	gtccgc当地	tggccc当地	caatggc当地	ctggccgccc	6540
tggccggcct	gctgaaactc	tggctcaccg	acgaccgc当地	cacggc当地	ttc当地	6600
ccacgatcct	c当地	gcaagatcg	aagagaagca	ggacgagctt	ggcaagg	6660
tgatgggcgt	ggtccgc当地	agggcagagc	catgactttt	ttagccgcta	aaacggccgg	6720
ggggtgc当地	tgattgcca	gcacgtcccc	atgc当地	tcaagaagag	cgacttc当地	6780
gagctggat	tc当地	caagattcgg	aatacca	act	acgagaagga	6840
gtctacggga	ccgacttcat	tgccgataag	gtggattatc	tggacaccaa	ggcaccag	6900
gggtcaaaatc	aggaataagg	gcacattgc	ccggcgtgag	tggggcaat	cccgcaagga	6960
gggtgaatga	atcgacg	tgaccgaa	gcatacaggc	aagaactgat	cgacgc当地	7020
tttccgc当地	aggatgccc	aaccatcgca	agccgcacc	tcatgc当地	gccc当地	7080
accttccagt	ccgtc当地	gatggtccag	caagctacgg	ccaagatcg	gc当地	7140
gtgcaactgg	ctccccctgc	c当地	ccatcgcc	ccgtggagc	ttc当地	7200
ctcgaacagg	aggcggcagg	tttggc当地	tc当地	tc当地	aggaaadtatg	7260
acgaccaaga	agcgaaaaac	cgccggc当地	gacctggca	aacaggtc	cgaggcoa	7320
caggccgc当地	tgctgaaaca	cacgaagcag	cagatcaagg	aatgcag	ttc当地	7380
gatattgc当地	cgtggccg	cacgatgc	gc当地	acgacacggc	ccgctctgc	7440
ctgttacca	c当地	aaaaatccc	cgcgaggc	tgcaaaaca	ggtc当地	7500
cacgtcaaca	aggacgtgaa	gatcacctac	accggcgtc	agctgc当地	cgacgatgac	7560
gaactgggt	ggc当地	ggc当地	gcaagc	ccc当地	cgagccgatc	7620
accttca	ctacgag	ttgc当地	ctggc当地	cgatcaatgg	ccggta	7680
acgaaggcc	aggaatgc	gtc当地	c当地	c当地	c当地	7740
c当地	accttggaa	ggtgtcg	ctgca	c当地	c当地	7800
aagaaaaacgt	ccc当地	ggtc当地	gatc	gacgaggaaa	tc当地	7860

gaccactaca cgaaattcat atgggagaag taccgcaagg tgtcggcgc gggccgacgg 7920  
 atgttcgact atttcagctc gcaccgggag ccgtacccgc tcaagctgg aaccttccgc 7980  
 ctcatgtgcg gatcgattc cacccgcgtg aagaagtggc gcgagcaggt cggcgaagcc 8040  
 tgcgaagagt tgcgaggcag cggcctggt gaacacgcct gggtaatga tgacctggtg 8100  
 cattgcaaac gctagggcct tgggggtca gttccggctg ggggttcagc agccagcgct 8160  
 ttactggcat ttcaggaaca agcgggact gctcgacgca cttgcttcgc tcagtatcgc 8220  
 tcgggacgca cggcgcgc tacaactgc cgataaacag aggattaaaa ttgacaattg 8280  
 tgattaaggc tcagattcga cggcttggag cggccgacgt gcaggatttc cgcgagatcc 8340  
 gattgtcgcc cctgaagaaa gctccagaga tgggggtc cgtttacgag cacgaggaga 8400  
 aaaagccccat ggaggcggtc gctgaacggt tgcgagatgc cgtggcattc ggccctaca 8460  
 tcgacggcga gatcattggg ctgtcggtct tcaaaccaggaa ggacggcccc aaggacgctc 8520  
 acaaggcgca tctgtccggc gtttctgtgg agccgaaca gcgaggccga ggggtcgccg 8580  
 gtatgctgct gccccgggtt ccggccgggtt tattgctcgat gatgatcgctc cgacagattc 8640  
 caacgggaat ctggggatg cgcatttca tccctggcgc acttaatatt tcgctattct 8700  
 ggagcttggtt gtttatttcg gtctaccggc tgccggggcg ggtcgccgc acggtaggcg 8760  
 ctgtgcagcc gctgatggtc gtgttcatct ctggcgtct gctaggttagc ccgatacgat 8820  
 tggatggcggtt cctggggctt atttgcggaa ctgcggcggtt ggcgctgttg gtgttgcac 8880  
 caaacgcagc gctagatcct gtcggcgctc cagcggccct ggccggggcg gtttccatgg 8940  
 cggttcggaaac cgtgctgacc cgcaagtggc aacctcccgt gcctctgc acctttaccg 9000  
 cctggcaact ggccggccggaa ggacttctgc tcggtccagt agcttttagtg tttgatccgc 9060  
 caatcccgat gcctacaggaa accaatgttc tccggctggc gtggctccgc ctgatccgg 9120  
 cgggttaac ctacttcctt tgggtccggg ggtatctcgatc actcgaacct acagttttt 9180  
 ctttactggg ctttctcagc ccggggaccg ccgtgttgctt aggtgggttgc ttcttggatc 9240  
 agacgctgag tgcgttcaa atcatcgcc tcctgctcgat gatcggagttt atctggctgg 9300  
 gccaacgttc caaccgcact cctaggcgcc gtagatcttgc ccggaaatcg cttgaccgg 9360  
 catggcatag gcctatcgat tccacgatca gcatcgatc cttggccctg cggccgttcca 9420  
 aagcccgccg cgcaggcccg gcaggccagag caagtagagg gcagcgccctg caatccatgc 9480  
 ccacccgttc cacgttgtta tagaagccgc atagatcgcc gtgaagagga ggggtccgac 9540  
 gatcgaggtc aggctgggtga gcgccgcccag tgagccttgc agctgcccct gacgttccctc 9600  
 atccacactgc ctggacaaca ttgcttgcag cggccggcatt ccgatgccac ccgaagcaag 9660

caggaccatg atcggaaacg ccatccatcc ccgtgtcgcg aaggcaagca ggatgttagcc 9720  
 tgtgccgtcg gcaatcattc cgagcatgag tgcccgctt tcgcccggcc gggcggtac 9780  
 agggccggtg atcattgcct gggcgagtga atgcagaatg ccaaatacgccg caagcgaaat 9840  
 gccgatcgtg gtcgcgtccc agtggaaagcg atcctcgccg aaaatgaccc aaagcgcggc 9900  
 cggcacctgt cggacaaggta gcatgatgaa gaagaccgcc atcagggccg cgacgacggt 9960  
 catgccccgg gcccacccgga acgaagcgag cgggttgaga gcctcccgcc gtaacggccg 10020  
 gcgttcgcct ttgtgcgact cggcaaaaag gaaacagccc gtcaggaaat tgaggccgtt 10080  
 caaggctgcc gcggcgaaga acggagcgtg gggggagaaaa ccgccccatca gcccacccgag 10140  
 cacaggtccc gcgaccatcc cgaacccgaa acaggcgctc atgaagccga agtgcgcgc 10200  
 gcgcctcatcg ccatcagtga tatcgcaat ataagcgccg gctaccgccc cagtcgcccc 10260  
 ggtgatgccg gccacgatcc gtccgatata gagaacccaa aggaaaggcg ctgtcgccat 10320  
 gatggcgttag tcgacagtgg cgccggccag cgagacgagc aagattggcc gccgcccggaa 10380  
 acgatccgac agcgcgcccc gcacaggtgc gcaggcaaat tgcaccaacg catacagcgc 10440  
 cagcagaatg ccatagtggg cggtgacgtc gttcgagtga accagatcgc gcaggaggcc 10500  
 cggcagcacc ggcataatca ggccgatgcc gacagcgctg agcgcgacag tgctcagaat 10560  
 tacgatcagg ggtatgttgg gtttcacgtc tggcctccgg accagcctcc gctggtccga 10620  
 ttgaacgcgc ggattctta tcactgataa gttggggac atattatgtt tatcagtgtat 10680  
 aaagtgtcaa gcatgacaaa gttgcagccg aatacagtga tccgtgccgc cctggacctg 10740  
 ttgaacgagg tcggcgtaga cggctctgac acacgcaaac tggcggaaacg gttgggggtt 10800  
 cagcagccgg cgctttactg gcacttcagg aacaagcggg cgctgctcga cgcactggcc 10860  
 gaagccatgc tggcggagaa tcatacgcatt tcggtgccga gagccgacga cgactggcgc 10920  
 tcatttctga tcgggaaatgc ccgcagcttc aggcaaggcgc tgctcgcccta ccgcgatggc 10980  
 gcgcgcatcc atgcccggcac gcgaccgggc gcaccgcaga tggaaacggc cgacgcccgc 11040  
 ttgcgttcc tctgcgaggc gggttttcg gcccgggacg cggtaatgc gctgtgacata 11100  
 atcagctact tcactgttgg ggccgtgctt gaggagcagg cggcgacag cgatgcccgc 11160  
 gagcgcggcg gcaccgttga acaggctccg ctctcgccgc tggcggggc cgcgatagac 11220  
 gccttcgacg aagccggtcc ggacgcgcgc ttgcgaggcagg gactcgccgt gattgtcgat 11280  
 ggattggcga aaaggaggct cgttgcagg aacgttgaag gaccgagaaaa gggtgacgat 11340  
 tgatacagag ccgggtttgt cacccgtata agctgaagca ggcacaaatc agggaaataa 11400  
 acaaaaatccc gcatccccgg ataaagaaaa atcagggaaat taatggcctg atggattcc 11460

cgtggcgtcg aaagtgcggc accataggta tcagtcaccg cgatgagatc cttaccatt 11520  
 ccagagtctg gcggttgatt attaatttgc tgatatagag ctcagcccg ctggc 11575

<210> 6

<211> 1632

<212> DNA

<213> *Mentha spicata*

<400> 6  
 attagacat ccggaaacta caacccttct cgttggatg tcaacttcat ccaatcgott 60  
 ctcagtgact ataaggagga caaacacgtg attaggcctt ctgagctggt cactttggtg 120  
 aagatggAAC tggagaaaga aacggatcaa attcgacaac ttgagttgat cgatgacttg 180  
 cagaggatgg ggctgtccga tcatttccaa aatgagttca aagaaatctt gtcctotata 240  
 tatctcgacc atcactatta caagaaccct tttccaaaag aagaaaggga tctctactcc 300  
 acatctcttg catttaggct cctcagagaa catggtttc aagtcgcaca agaggtattc 360  
 gatagtttca agaacgagga gggtgagttc aaagaaagcc tttagcgcacga caccagagga 420  
 ttgttgcaac tgtatgaagc ttcccttctg ttgacggaag gcaaaccac gctcgagtca 480  
 gcgaggaaat tcgcccaccaa atttttggag gaaaaagtga acgagggtgg tggatggc 540  
 gaccttttaa caagaatcgc atattctttg gacatccctc ttcatggag gattaaaagg 600  
 ccaaatgcac ctgtgtggat cgaatggtat aggaagaggc ccgcacatgaa tccagtagtg 660  
 ttggagcttg ccatactcga cttaaatatt gttcaagcac aatttcaaga agagctaaa 720  
 gaatccttca ggtggggag aaatactggg tttgttgaga agctgccctt cgcaaggat 780  
 agactggtg aatgctactt ttgaaatact gggatcatcg agccacgtca gcatgcaagt 840  
 gcaaggataa tgatggcaaa agtcaacgct ctgattacgg tgatcgatga tattttatgat 900  
 gtctatggca ctttagaaga actcgaacaa ttcaactgacc tcattcgaag atgggatata 960  
 aactcaatcg accaacttcc cgattacatg caactgtgct ttcttgact caacaacttc 1020  
 gtcgatgata catcgatcga ttttatgaag gagaaaggcg tcaacgttat accctacctg 1080  
 cggcaatcg gggttgattt ggcggataag tatatggtag aggcacggtg gttctacggc 1140  
 gggcacaaac caagtttgg aagttttt gagaactcat ggcagtcgat aagtggccc 1200  
 tgtatgttaa cgcacatatt cttccgagta acagattcgt tcacaaagga gaccgtcgac 1260  
 agtttgcata aataccacga tttagttcgt tggtcaccc tcgttctgac gtttgctgat 1320

gatttggaa cctcggtgga agaggtgagc agagggatg tgccgaaatc acttcagtgc 1380  
 tacatgagtg actacaatgc atcgaggcg gaggcgcgga agcacgtgaa atggctgata 1440  
 gcggaggtgt ggaagaagat gaatgcggag agggtgtcga aggattctcc attcggcaaa 1500  
 gattttatag gatgtgcagt tgattttagga aggatggcgc agttgatgta ccataatgga 1560  
 gatgggcacg gcacacaaca ccctattata catcaacaaa tgaccagaac cttattcag 1620  
 ccctttgcat ga 1632

<210> 7

<211> 543

<212> PRT

<213> *Mentha spicata*

<400> 7

Met Arg Arg Ser Gly Asn Tyr Asn Pro Ser Arg Trp Asp Val Asn Phe  
 1 5 10 15

Ile Gln Ser Leu Leu Ser Asp Tyr Lys Glu Asp Lys His Val Ile Arg  
 20 25 30

Ala Ser Glu Leu Val Thr Leu Val Lys Met Glu Leu Glu Lys Glu Thr  
 35 40 45

Asp Gln Ile Arg Gln Leu Glu Leu Ile Asp Asp Leu Gln Arg Met Gly  
 50 55 60

Leu Ser Asp His Phe Gln Asn Glu Phe Lys Glu Ile Leu Ser Ser Ile  
 65 70 75 80

Tyr Leu Asp His His Tyr Tyr Lys Asn Pro Phe Pro Lys Glu Glu Arg  
 85 90 95

Asp Leu Tyr Ser Thr Ser Leu Ala Phe Arg Leu Leu Arg Glu His Gly  
 100 105 110

Phe Gln Val Ala Gln Glu Val Phe Asp Ser Phe Lys Asn Glu Glu Gly  
 115 120 125

Glu Phe Lys Glu Ser Leu Ser Asp Asp Thr Arg Gly Leu Leu Gln Leu  
 130 135 140

Tyr Glu Ala Ser Phe Leu Leu Thr Glu Gly Glu Thr Thr Leu Glu Ser  
 145 150 155 160

Ala Arg Glu Phe Ala Thr Lys Phe Leu Glu Glu Lys Val Asn Glu Gly  
 165 170 175

Gly Val Asp Gly Asp Leu Leu Thr Arg Ile Ala Tyr Ser Leu Asp Ile  
 180 185 190

Pro Leu His Trp Arg Ile Lys Arg Pro Asn Ala Pro Val Trp Ile Glu  
 195 200 205

Trp Tyr Arg Lys Arg Pro Asp Met Asn Pro Val Val Leu Glu Leu Ala  
 210 215 220

Ile Leu Asp Leu Asn Ile Val Gln Ala Gln Phe Gln Glu Glu Leu Lys  
 225 230 235 240

Glu Ser Phe Arg Trp Trp Arg Asn Thr Gly Phe Val Glu Lys Leu Pro  
 245 250 255

Phe Ala Arg Asp Arg Leu Val Glu Cys Tyr Phe Trp Asn Thr Gly Ile  
 260 265 270

Ile Glu Pro Arg Gln His Ala Ser Ala Arg Ile Met Met Gly Lys Val  
 275 280 285

Asn Ala Leu Ile Thr Val Ile Asp Asp Ile Tyr Asp Val Tyr Gly Thr  
 290 295 300

Leu Glu Glu Leu Glu Gln Phe Thr Asp Leu Ile Arg Arg Trp Asp Ile  
 305 310 315 320

Asn Ser Ile Asp Gln Leu Pro Asp Tyr Met Gln Leu Cys Phe Leu Ala  
 325 330 335

Leu Asn Asn Phe Val Asp Asp Thr Ser Tyr Asp Val Met Lys Glu Lys  
 340 345 350

Gly Val Asn Val Ile Pro Tyr Leu Arg Gln Ser Trp Val Asp Leu Ala  
 355 360 365

Asp Lys Tyr Met Val Glu Ala Arg Trp Phe Tyr Gly Gly His Lys Pro  
 370 375 380

Ser Leu Glu Glu Tyr Leu Glu Asn Ser Trp Gln Ser Ile Ser Gly Pro  
 385 390 395 400

Cys Met Leu Thr His Ile Phe Phe Arg Val Thr Asp Ser Phe Thr Lys  
 405 410 415

Glu Thr Val Asp Ser Leu Tyr Lys Tyr His Asp Leu Val Arg Trp Ser  
 420 425 430

Ser Phe Val Leu Arg Leu Ala Asp Asp Leu Gly Thr Ser Val Glu Glu  
 435 440 445

Val Ser Arg Gly Asp Val Pro Lys Ser Leu Gln Cys Tyr Met Ser Asp  
 450 455 460

Tyr Asn Ala Ser Glu Ala Glu Ala Arg Lys His Val Lys Trp Leu Ile  
 465 470 475 480

Ala Glu Val Trp Lys Lys Met Asn Ala Glu Arg Val Ser Lys Asp Ser  
 485 490 495

Pro Phe Gly Lys Asp Phe Ile Gly Cys Ala Val Asp Leu Gly Arg Met  
500 505 510

Ala Gln Leu Met Tyr His Asn Gly Asp Gly His Gly Thr Gln His Pro  
515 520 525

Ile Ile His Gln Gln Met Thr Arg Thr Leu Phe Glu Pro Phe Ala  
530 535 540